(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 April 2004 (01.04.2004)

PCT

(10) International Publication Number
WO 2004/027088 A3

(51) International Patent Classification[7]: C12Q 1/68, C12N 9/10

(21) International Application Number: PCT/CA2003/001269

(22) International Filing Date: 20 August 2003 (20.08.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/412,002 20 September 2002 (20.09.2002) US

(71) Applicant *(for all designated States except US)*: **UNIVERSITÉ LAVAL** [CA/CA]; Cité universitaire, Québec, Québec G1K 7P4 (CA).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **GUILLEMETTE, Chantal** [CA/CA]; 7605 La Fraîcheur, Québec, Québec G2K 2C7 (CA).

(74) Agent: **OGILVY RENAULT**; Suite 1600, 1981 McGill College Avenue, Montréal, Québec H3A 2Y3 (CA).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report: 19 August 2004

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** METHOD FOR DETERMINING THE PREDISPOSITION OF PATIENTS TO TOXICITY OR LACK OF EFFICACY OF A DRUG

(57) **Abstract:** The present invention relates to a method for determining predisposition to a physiological reaction in a patient. Particularly, the present invention relates to a method for determining a predisposition to toxicity induced by a camptothecin analog or to an immunosuppressive mycophenolic acid-based therapy. This method comprises the characterization of nucleic acid sequences from the patient. The nucleic acid sequence encodes for an amino acid sequence or regulates the expression of UGT1A1, UGT1A7, UGT1A9 or their polymorphic variants. The method also comprises the analysis of haplotypic variation within these genes.

WO 2004/027088 A3

International[redacted]plication No
PCT/CA 03/01269

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68 C12N9/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12Q C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 395 481 B1 (RATAIN MARK J ET AL) 28 May 2002 (2002-05-28) examples 1-7; table 1 column 6, line 54 - column 7, line 9 claims 9-26,43-63	1-15,23, 24
X	WO 02/48400 A (ANDO YU-UICHI ;SHIMOKATA KAORU (JP); HASEGAWA YOSHINORI (JP); NAGO) 20 June 2002 (2002-06-20) abstract; figure 4 claims 1-31 & EP 1 352 970 A (NAGOYA IND SCIENCE RES I) 15 October 2003 (2003-10-15)	1-15,23, 24

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 March 2004	19.04.04

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hermann, P

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GAGNE J-F ET AL: "COMMON HUMAN UGT1A POLYMORPHISMS AND THE ALTERED METABOLISM OF IRINOTECAN ACTIVE METABOLITE 7-ETHYL-10-HYDROXYCAMPTOTHECIN (SN-38)" MOLECULAR PHARMACOLOGY, BALTIMORE, MD, US, vol. 62, no. 3, September 2002 (2002-09), pages 608-617, XP009001388 ISSN: 0026-895X Experimental procedures Results Discussion -----	1-15,23, 24
X	ANDO Y ET AL: "Polymorphisms of UDP-glucuronosyltransferase and pharmacokinetics of irinotecan" THERAPEUTIC DRUG MONITORING, NEW YORK, NY, US, vol. 24, no. 1, February 2002 (2002-02), pages 111-116, XP002967773 ISSN: 0163-4356 the whole document -----	1-15,23, 24
X	INNOCENTI F ET AL: "PHARMACOGENETICS OF ANTICANCER AGENTS: LESSONS FROM AMONAFIDE AND IRINOTECAN" DRUG METABOLISM AND DISPOSITION, WILLIAMS AND WILKINS., BALTIMORE, MD, US, vol. 29, no. 4, PART 2, April 2001 (2001-04), pages 596-600, XP001008560 ISSN: 0090-9556 the whole document -----	1-15,23, 24
X	ANDO Y ET AL: "POLYMORPHISMS OF UDP-GLUCURONOSYLTRANSFERASE GENE AND IRINOTECAN TOXICITY: A PHARMACOGENETIC ANALYSIS" CANCER RESEARCH, AMERICAN ASSOCIATION FOR CANCER RESEARCH, BALTIMORE, MD, US, vol. 60, 15 December 2000 (2000-12-15), pages 6921-6926, XP002909302 ISSN: 0008-5472 the whole document -----	1-15,23, 24
X	ANDO Y ET AL: "UGT1A1 GENOTYPES AND GLUCURONIDATION OF SN-38, THE ACTIVE METABOLITE OF IRINOTECAN" ANNALS OF ONCOLOGY, KLUWER, DORDRECHT, NL, vol. 9, 1998, pages 845-847, XP002909100 ISSN: 0923-7534 the whole document ----- -/--	1-15,23, 24

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RAMIREZ J ET AL: "IN VITRO CHARACTERIZATION OF HEPATIC FLAVOPIRIDOL METABOLISM USING HUMAN LIVER MICROSOMES AND RECOMBINANT UGT ENZYMES" PHARMACEUTICAL RESEARCH, NEW YORK, NY, US, vol. 19, no. 5, May 2002 (2002-05), pages 588-594, XP009019622 ISSN: 0724-8741 the whole document -----	1-20, 25-29
X	VOGEL A ET AL: "GENETIC LINK OF HEPATOCELLULAR CARCINOMA WITH POLYMORPHISMS OF THE UDP-GLUCURONOSYLTRANSFERASE UGT147 GENE" GASTROENTEROLOGY, W.B.SAUNDERS COMPANY, PHILADELPHIA, US, vol. 121, no. 5, November 2001 (2001-11), pages 1136-1144, XP001119046 ISSN: 0016-5085 the whole document -----	1-15,21, 22,25,26
X	US 2002/016293 A1 (RATAIN MARK J ET AL) 7 February 2002 (2002-02-07) paragraph ‘0070! - paragraph ‘0073!; claims 77-83 -----	1-20, 25-29
X	GUILLEMETTE C ET AL: "STRUCTURAL HETEROGENEITY AT THE UDP-GLUCURONOSYLTRANSFERASE 1 LOCUS: FUNCTIONAL CONSEQUENCES OF THREE NOVEL MISSENSE MUTATIONS IN THE HUMAN UGT1A7 GENE" PHARMACOGENETICS, CHAPMAN & HALL, LONDON, GB, vol. 10, no. 7, October 2000 (2000-10), pages 629-644, XP009002994 ISSN: 0960-314X the whole document -----	1-15,21, 22,25,26
X	ANDO MAKI ET AL: "Genetic polymorphisms of the UDP-glucuronosyltransferase 1A7 gene and irinotecan toxicity in Japanese cancer patients." JAPANESE JOURNAL OF CANCER RESEARCH : GANN. MAY 2002, vol. 93, no. 5, May 2002 (2002-05), pages 591-597, XP002273391 ISSN: 0910-5050 the whole document -----	1-15,21, 22,25,26

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TUKEY R H ET AL: "HUMAN UDP-GLUCURONOSYLTRANSFERASES: METABOLISM, EXPRESSION, AND DISEASE" ANNUAL REVIEW OF PHARMACOLOGY AND TOXICOLOGY, ANNUAL REVIEW INC., PALO ALTO, CA, US, vol. 40, 2000, pages 581-616,2PAGES, XP009003008 ISSN: 0362-1642 the whole document -----	1-15,23, 24
A	MACKENZIE PETER I: "Identification of uridine diphosphate glucuronosyltransferases involved in the metabolism and clearance of mycophenolic acid" THERAPEUTIC DRUG MONITORING, vol. 22, no. 1, February 2000 (2000-02), pages 10-13, XP009019790 ISSN: 0163-4356 the whole document -----	11
A	WOOLLEY ADAM T ET AL: "Direct haplotyping of kilobase-size DNA using carbon nanotube probes" NATURE BIOTECHNOLOGY, NATURE PUBLISHING, US, vol. 18, no. 7, July 2000 (2000-07), pages 760-763, XP002234415 ISSN: 1087-0156 the whole document -----	1-15,21, 22,25,26

International application No.
PCT/CA 03/01269

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [ ] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. [X] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

As a result of the prior review under R. 40.2(e) PCT, no additional fees are to be refunded.

1. [X] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[X] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: -

Claims 25- 29 lack clarity (Article 6 PCT) due to the expression "fragment thereof" which does not have any commonly well accepted definition in the art. Moreover the description is silent as regards a definition of said expression, or any length characterising the said fragments. Therefore the search for claims 25-29 has been limited to nucleic acids and molecules having the SEQ. IDs. recited in the claims and the complements thereof.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.5), should the problems which led to the Article 17(2) declaration be overcome.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-15, 25 and 26 (all in part), 23 and 24 (in full)

   Invention 1 relates to a method for determining the predisposition of patients to toxicity or lack of efficacy of a biologically active compound comprising detecting polymorphism or haplotic variation in UGT1A1 gene wherein the presence of said polymorphism or haplotic variation is indicative of said predisposition.

   ---

2. claims: 1-15, 25 and 26 (all in part), 21 and 22 (in full)

   Invention 2 relates to i) a method for determining the predisposition of patients to toxicity or lack of efficacy of a biologically active compound comprising detecting polymorphism or haplotic variation in UGT1A7 gene wherein the presence of said polymorphism or haplotic variation is indicative of said predisposition; and ii) the isolated nucleotide sequence comprising sequences presenting said polymorphism or variation.

   ---

3. claims: 1-15, 25 and 26 (in part), 16-20 and 27-29 (in full)

   Invention 3 relates to i) a method for determining the predisposition of patients to toxicity or lack of efficacy of a biologically active compound comprising detecting polymorphism or haplotic variation in UGT1A9 gene wherein the presence of said polymorphism or haplotic variation is indicative of said predisposition; ii) the isolated nucleotide sequence comprising sequences presenting said polymorphism or variation; and iii) the isolated amino acid sequence comprising the translation of said polymorphism or variation.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/CA 03/01269

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
US 6395481	B1	28-05-2002	US	2002115097	A1	22-08-2002
WO 0248400	A	20-06-2002	AU	2111202	A	24-06-2002
			EP	1352970	A1	15-10-2003
			WO	0248400	A1	20-06-2002
EP 1352970	A	15-10-2003	AU	2111202	A	24-06-2002
			EP	1352970	A1	15-10-2003
			WO	0248400	A1	20-06-2002
US 2002016293	A1	07-02-2002	AU	5361801	A	07-11-2001
			WO	0180896	A2	01-11-2001

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 April 2004 (01.04.2004)

PCT

(10) International Publication Number
WO 2004/027088 A2

(51) International Patent Classification[7]: C12Q 1/68, C12N 9/10

(21) International Application Number: PCT/CA2003/001269

(22) International Filing Date: 20 August 2003 (20.08.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/412,002 20 September 2002 (20.09.2002) US

(71) Applicant *(for all designated States except US)*: UNIVERSITÉ LAVAL [CA/CA]; Cité universitaire, Québec, Québec G1K 7P4 (CA).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: GUILLEMETTE, Chantal [CA/CA]; 7605 La Fraîcheur, Québec, Québec G2K 2C7 (CA).

(74) Agent: OGILVY RENAULT; Suite 1600, 1981 McGill College Avenue, Montréal, Québec H3A 2Y3 (CA).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/027088 A2

(54) **Title:** METHOD FOR DETERMINING PREDISPOSITION TO A PHYSIOLOGICAL REACTION IN A PATIENT.

(57) **Abstract:** The present invention relates to a method for determining predisposition to a physiological reaction in a patient. Particularly, the present invention relates to a method for determining a predisposition to toxicity induced by a camptothecin analog or to an immunosuppressive mycophenolic acid-based therapy. This method comprises the characterization of nucleic acid sequences from the patient. The nucleic acid sequence encodes for an amino acid sequence or regulates the expression of UGT1A1, UGT1A7, UGT1A9 or their polymorphic variants. The method also comprises the analysis of haplotypic variation within these genes.

METHOD FOR DETERMINING PREDISPOSITION TO A PHYSIOLOGICAL REACTION IN A PATIENT

TECHNICAL FIELD

The present invention relates to a method for determining predisposition to a physiological reaction to a xenobiotic, a drug or an endogenously secreted compound, in a patient. Particularly, the present invention consists in a method comprising the characterization of a nucleic acid sequence from a patient. These nucleic acid sequences encode for amino acid sequences or regulate the expression of genes.

BACKGROUND ART

Recent evidences support the concept that polymorphic variation in genes encoding metabolism enzymes (MEs) are likely to play an important role in clinical response to therapeutic drugs and in exogenous or endogenous compound elimination.

Interindividual variations in response to a drug or to exogenous or endogenous compounds can be classified in three groups. The first segment of the population is known as poor metabolizers (PMs). These individuals often show accumulation of drugs or metabolites caused by a genetic defect in metabolizing enzymes and increased predisposition to adverse drug reactions is an important consequence of PM genotypes. In opposite, ultrarapid metabolizers (UMs) eliminate drugs excessively rapidly from the body. These patients, for example, do not develop sufficient high plasma levels of drugs and therefore do not respond to treatments, also giving rise to both clinical and economical complications. The remaining proportion of the population categorized as "normal" patients are named extensive metabolizers (EMs).

Some researchers have studied pharmacogenetics of human drug-metabolizing enzymes (DME), more specifically enzymes of the glucuronidation pathway and have demonstrated that glucuronidation, like other DME pathways, is also subject to interindividual variations. The glucuronidation reaction is catalyzed by UDP-glucuronosyltransferase enzymes (UGTs), a set of enzymes that

increase the polarity of xenobiotics, drugs and endogenous compounds to facilitate their excretion from the body. Glucuronidation reaction occurs on different functional groups that include hydroxyl, carboxyl, amino and sulfur. UGTs have the most important effect in both detoxification and promotion of excretion, via both urine and bile. Apart from being a major biochemical pathway well known for its role in drug metabolism, the glucuronidation system is also clearly involved in the homeostasis of numerous endogenous molecules, including steroids, thyroid hormones and bile acids.

Any perturbation in the glucuronidation pathway has the potential to modify the elimination, the detoxification or the pharmacokinetic parameters of a given drug, and consequently drug clearance. As a result, in situations where the activity of the glucuronidation pathway is reduced, it is to be expected that changes in the biological activity, sometimes toxicity, of the compounds will ensue. Therefore, the human genetic variations leading to differences in the glucuronidation rates could influence the activity of drugs and other chemicals, which undergo this conjugation.

As example, SN-38 or 7-ethyl-10-hydroxycamptothecin, which is the pharmacologically active metabolite of the anticancer drug irinotecan, undergoes extensive glucuronidation in human to form SN-38-G (10-O-glucuronyl-SN-38) and goes through significant biliary excretion and enterohepatic circulation. This drug is used globally in the first line treatment of advanced metastatic colorectal cancer (CRC). A major drawback of irinotecan-based chemotherapy is the high incidence of severe hematological and gastrointestinal toxicities, such as diarrhea. Diarrhea is believed to be secondary to the biliary excretion of SN-38, the extent of which is determined by SN-38 glucuronidation. Incidences of irinotecan-induced diarrhea can be serious and do not respond adequately to conventional antidiarrheal agents. It is believed that SN-38-G can be deconjugated to form SN-38 by intestinal glucuronidase enzyme, and further causes diarrhea by direct enteric injury. An inverse relationship between SN-38 glucuronidation rates and severity of diarrhea incidences in patients treated with irinotecan has been shown. These findings indicated that glucuronidation of SN-38 protects against irinotecan-

induced gastrointestinal toxicities. Therefore, the conversion of SN-38 to SN-38-G by both hepatic and intestinal UGTs is a critical step in the sequential metabolic pathway of irinotecan, and consequently in drug response and toxicity. Over the existing human UGTs, UGT1A1, UGT1A7 and UGT1A9 are known in the art to be SN-38 conjugators. On the other hand, UGT1A1 and UGT1A9 are highly expressed in the liver, the primary organ involved in the detoxification of irinotecan, and also in the gastrointestinal tract (GI) where toxicity takes place.

Mycophenolic acid (MPA) is also an extensively glucuronidated drug for which an interindividual variation of glucronidation rates is observed. MPA is a metabolite of mycophenolate mofetil (MMF), and is commonly used as immunosuppressive agent. As MPA is known to be conjugated exclusively by the liver UGT1A9, interindividual variation observed with this substrate is therefore attributable only to the UGT1A9 isoform. The study of UGT1A9 polymorphic variations thus plays a critical role in the control of immunosuppressive therapies and management of graft rejection.

Genetic variations among UGT isoforms have been demonstrated to be also implicated in the interindividual physiological response to drug administration. Therefore, glucuronidation pathway represented a target for many groups as a way to control irinotecan-associated side effects.

As example, international patent publication number WO 96/01127 describes a method and pharmaceutical compositions to reduce side effects of camptothecin analogs such as irinotecan, therefore reducing associated side effects. This reduction of toxicity would occur by reducing biliary transport or increasing UGT activity, by administrating concomitantly a transport inhibitor or an UGT inducer.

US Patent no. 6,395,481 reports a method for detecting TA repeats polymorphic variations within the promoter region of the *UGT1A1* gene to evaluate predispositions to drug sensitivity associated with low levels of UGT enzymes expression.

International patent publication number WO 02/48400 reports a method for estimating the susceptibility in an individual to adverse side effects caused by the administration of irinotecan. This method is also based on the evaluation of the TA repeats within the promoter region of *UGT1A1*, but also includes the analysis of single nucleotide polymorphisms at two other positions within the exon 1.

International patent publication number WO 03/013536 reports a method for selecting a suitable iriontecan therapy for a cancer patient that comprises determining whether the patient has one or multiple variant alleles of the UGT1A1 gene and adjusting irinotecan dosage and/or UGT1A1-modulating drugs consequently.

Considering overlapping substrate specificities of UGT enzymes, it is noteworthy that a higher expression of UGT1A1 protein resulting from an increased gene expression could complement a deficient glucuronidation activity of an altered UGT1A9 protein, or the contrary. An individual harboring two mutated genotypes would therefore have a normal phenotype and is less susceptible to develop a toxicity to a drug than a patient having the low metabolizer phenotype. Therefore, the genotyping studies that consider only one gene encoding a xenobiotic conjugating enzyme are less likely to be accurate than a global analysis of the whole set of genes.

Based on the state of the prior art described hereinabove, it would be highly desirable to be provided with a new diagnostic tool to determine accurately a predisposition to physiological adverse response following drug administration in standard conditions. This would allow to provide physicians with guiding means in determining drugs to be used in a specific treatment.

DISCLOSURE OF INVENTION

One aim of the present invention is to provide a method for determining a predisposition to a physiological reaction of an individual to a biologically active compound. This method comprises characterizing nucleotide sequence of the individual for at least one of the *UGT1A1*, *UGT1A7* or *UGT1A9* gene, or a pert

thereof. The presence of at least one polymorphic or haplotypic variation in this nucleotide sequence is indicative of the predisposition to the physiological reaction.

In accordance with the method described herein, the predisposition may be a hereditary predisposition and the physiological reaction in the patient may be a beneficial reaction, an adverse reaction or a side effect to a compound.

Another aim of the present invention is to provide a method wherein determining the genetic sequence comprises determining the presence of at least one polymorphic or haplotypic variation in *UGT1A1*, *UGT1A7* or *UGT1A9* gene. These variations may include variations of the number of TA repeats in a TATA box of the UGT1A1 gene, $C^{-2208}T$ substitution, $C^{-2152}T$ substitution, $C^{-2141}T$ substitution, $T^{-1887}G$ substitution, $T^{-1818}C$ substitution, $C^{-665}T$ substitution, $T^{-440}C$ substitution, $C^{-331}T$ substitution, $T^{-275}A$ substitution, $G^{-87}A$ substitution, $G^{8}A$ missence mutation, a $T^{98}C$ missence mutation, or a combination of these variations in the UGT1A9 gene. Alternatively and/or additionally, $G^{353}T$, $T^{397}G$, $C^{401}A$, $G^{402}A$, $G^{427}C$ or $T^{632}C$ missense mutations can be determined in the *UGT1A7* gene.

Another aim of the present invention is to provide a nucleotide sequence for determining a predisposition to a physiological reaction comprising at least one nucleotide sequence selected from the group consisting of SEQ ID NO: 36 to SEQ ID NO: 68, or the complementary sequences thereof.

For the purpose of the present invention the following terms are defined below.

The expression "adverse physiological reaction" is intended to mean any physiological reaction that provides a negative physiological effect to an individual. .

The term "ASO" is intended to mean Allele Specific Oligonucleotide analysis.

The term "ASP" is intended to mean Allele Specific PCR analysis.

The expressions "beneficial physiological reaction" or "beneficial reaction" are intended to mean any physiological reaction that provides a positive physiological effect to an individual.

The term "BPD" is intended to mean benzo(a)pyrene-trans-7,8-dihydrodiol.

The term "CPT-11" is intended to mean 7-ethyl-10-[4-(1-piperidino)-1-piperidino] carbonyloxy camptothecin.

The term "DHPLC" is intended to mean denaturing high-performance liquid chromatography.

The term "gene" is intended to mean a segment of nucleic acid involved in producing a polypeptide chain; it includes regions preceding the coding region (promoter, leader sequence), regions following coding region (trailer) and intervening sequences (introns) between individual coding segments (exons).

The term "GI" is intended to mean gastrointestinal tract.

The term "MPA" us intended to mean mycophenolic acid.

The term "PhIP" is intended to mean 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine.

The term "RFLP" is intended to mean Restriction Fragment Length Polymorphism analysis.

The term "SN-38" is intended to mean 7-ethyl-10-hydroxycamptothecin.

The term "SSCP" is intended to mean Single Strand Conformation Polymorphism analysis.

The term "UGT" is intended to mean uridine diphospho-glucuronosyltransferase.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates the metabolic pathway of irinotecan hydroclorine (CPT-11);

Fig. 2 illustrates the entero-hepatic cycle of irinotecan biotransformation;

Fig. 3 illustrates the major role of UGT1A9 in SN-38 glucuronidation;

Fig. 4 illustrates the distribution of SN-38-G formation by human liver samples.

Figs. 5a to 5f illustrate methods for detecting SNPs;

Figs. 6a to 6d illustrate the missence mutations in the human first exons of *UGT1A7* and *UGT1A9* genes;

Fig. 7 illustrates the expression of the UGT1A9 and UGT1A9 proteins in human liver microsomes;

Figs. 8a to 8e illustrate the effect of UGT1A9 promoter polymorphisms on protein expression;

Figs. 9 illustrates the effect of the UGT1A9 (-2152) polymorphic variation on MPA glucuronidation activity;

Fig. 10 illustrates the effect of the UGT1A9 (-1818) polymorphic variation on SN-38 glucuronidation activity;

Figs. 11a to 11d illustrate the effect of the UGT1A9 (-665) polymorphic variation on glucuronidation activity;

Figs 12 illustrates the effect of UGT1A9 (-275) polymorphic variation on MPA glucuronidation activity;

Figs. 13a and 13b illustrate the correlation between the UGT1A9 protein expression and glucuronidation activity;

Figs. 14a to 14d illustrate the relative expression of UGT1A7 and UGT1A9 protein and their relative activities on SN-38;

Figs. 15a to 15c illustrate the glucuronidation rates of the variant UGT1A9 allozymes;

Figs. 16a to 16l illustrate the immunofluorescence localization of UGT1A9*1, UGT1A9*2 and UGT1A9*3;

Figs. 17a to 17c illustrate the relationship between UGT1A1 TATA box polymorphic variations and protein expression or glucuronidation activity;

Figs. 18a and 18b illustrate the correlative association between UGT1A1 protein expression and glucuronidation activity; and

Fig. 19 illustrates the predictive value of the haplotype determination of UGT1A9 and UGT1A1; and

Figs. 20a and 20b illustrate a sequence alignment of UGT1A proteins at selected positions.

MODES OF CARRYING OUT THE INVENTION

In accordance with the present invention, there is provided a method for determining a predisposition to a physiological reaction in an individual comprising characterizing nucleotide sequence of at least one of the *UGT1A1*, *UGT1A7* or *UGT1A9* gene or a part thereof of the individual, where the nucleotide sequence is indicative of the predisposition to a physiological reaction. The individual of the present invention is a human or an animal, but is preferably a patient having a colorectal cancer or a solid tumor. The predisposition determined with the present method is any higher or lower susceptibility, sensibility, diathesis, proneness, proclivity, tendency, sensitivity, responsiveness, resistance or constitutional sickness to the physiological reaction. This predisposition may be a hereditary predisposition, a non-hereditary congenital predisposition or an acquired predisposition.

The physiological reaction of the present invention comprises a beneficial reaction to a compound, an adverse reaction to a compound or a side effect. Among predisposition to an adverse physiological reaction to a compound, toxicity induced by an anti-cancer drug or a decreased responsiveness to an immunosuppressive agent are preferred. Toxicity to drug may be caused by an increased concentration of the drug in plasma, this increased concentration

being attributable to a lower glucuronidation metabolism of this compound or a decreased responsiveness to a drug, the latter being induced by an excessive glucuronidation-mediated elimination form of this compound from the organism. An anti-cancer agent that can be targeted through carrying out the present invention can be a camptothecin analog, such as 7-ethyl-10-[4-(1-piperidino)-1-piperidino] carbonyloxy camptothecin (irinotecan, CPT-11) or 7-ethyl-10-hydroxycamptothecin (SN-38).

As CPT-11 or its active metabolite SN-38 are topoisomerase inhibitors, cells showing higher levels of these enzymes are likely more sensitive to topoisomerase inhibition. Resistance to the drug occurs generally in cells that have low levels of topoisomerase. Resistance to irinotecan may also result from reduced conversion of the inactive prodrug CPT-11 to SN-38, attributable to reduced enzyme levels or, possibly, enzyme mutations. Additionally, an increased catabolic processing of the inhibitors contributes to reduce their availability within the cell, lowers inhibitor activity and favors drug resistance. It has also been reported that human colon tumors express high levels of the multiple-drug-resistance (MDR1) proteins. This class of enzyme may limit access of certain drugs to cells. *In vitro* data have demonstrated that camptothecin and its noncharged derivatives such as irinotecan overcome MDR1-mediated resistance. MDR1-mediated resistance to irinotecan may result from its rapid passive diffusion, its absence of interaction with MDR1, or a combination of both characteristics.

Alternatively, the sensitivity to drugs, as for example anticancer drugs, can be observed in cell lines deficient in DNA repair mechanisms. Indeed, DNA repair mechanisms can reverse drug-induced damage caused to the DNA. Therefore, DNA damage that goes unrepaired may result in significant genetic alterations or apoptosis.

The adverse physiological reaction as intended herein does not include the side effects observed with the majority of the population treated with the drug, but comprises physiological reactions that cause more serious threats in particular patients than what is generally expected with that drug in a majority of patients.

In fact, the susceptibility, sensitivity, responsiveness or resistance is higher or lower to what is observed in a patient having an anticipated physiological reaction to the same drug. These adverse physiological reactions are generally traduced by gastrointestinal, hematologic, hepatic, dermatologic, respiratory and neurologic disorders. Although gastrointestinal adverse reactions include nausea and vomiting, the most preoccupying and severe side effect observed is diarrhea. It has been observed that this particular toxicity is attributable to an accumulation of unconjugated SN-38 in the intestine. As SN-38 metabolization rates inversely correlate with the intensity of diarrhea in patients treated with increasing doses of CPT-11, the interindividual differences in pharmacokinetics of SN-38 are suggested to be responsible for the variation in drug side effects. Glucuronidation, which participates in the catabolic process of SN-38 is thus proposed to participate to this interindividual variation and the UGT1A9 enzyme would be responsible, at least in part, for these glucuronidation variations. As example, the UGT1A9(C^3Y) and UGT1A9($M^{33}T$) isoforms, trivially named UGT1A9*2 and UGT1A9*3, respectively, were shown to have a significantly reduced glucuronidation efficiency toward SN-38 (see Table 1). Therefore, individuals that hold one of these polymorphic variations would be more susceptible to present such adverse physiological reactions.

A person skilled in the art will understand that the invention is not limited to adverse physiological reactions to camptothecin analogs but rather finds uses in the determination of predisposition to physiological reactions to any other glucuronidated compound. Clinically and toxicologically important compounds include mycophenolic acid (MPA), flavopiridol, an anticancer agent under development and a number of xenobiotics, particularly a variety of pre-carcinogens such as the benzo(a)pyrene-trans-7,8-dihydrodiol (BPD), precursor to the potent mutagen benzo(a)pyrene-7,8-dihydrodiol-9,10-epoxide. Glucuronidation is an effective transforming pathway of pyrene to the 1-pyrenylglucuronide, a well-known urinary biomarker for the assessment of human exposure to polycyclic aromatic hydrocarbons. In addition, some UGT isoforms, such as UGT1A9, play a critical role in the detoxification of food-borne carcinogenic heterocyclic amines. Among those, 2-amino-1-methyl-6-

phenylimidazo[4,5-b]pyridine (PhIP), the most abundant carcinogenic heterocyclic amine found in well-cooked meats, has been shown to be extensively glucuronidated by UGT1A9 in humans. Genetic polymorphisms is a possible determinant factor of detoxifying UGT1A9 activity and the large interindividual variability in the metabolism of these carcinogens and therapeutics drugs. Finally, a skilled artisan will understand that the present invention also concerns endogenously produced compounds that include, but are not limited to steroids, hormones, fatty acids or bilirubin.

The method of the present invention may further comprise a step of obtaining a nucleic acid sample from the individual and/or extracting nucleic acid material from the biological sample. In such cases, the nature of the biological sample may be adapted for the purpose of the determination and may include saliva, semen, blood, hairs or any specimen comprising at least one cell from a human origin. This specimen can be collected directly on a human body or, alternatively, on any object on which nucleic acid molecules from a human origin could be found. The latter option is of particular interest in cases where inter-generation transmission of a gene (pedigree) is investigated, some members of the cohorts having disappeared. Nucleic acid extraction may include a further step of amplification to ensure an appropriate availability of material, wherein said amplification is preferably performed by polymerase chain reaction (PCR) amplification, wherein PCR amplification is performed using primers that specifically hybridize to a UGT1A9-encoding nucleic acid sequence. Nucleic acid molecules can be either single strand (ss) or double strand (ds) RNA or DNA, as well as DNA/RNA hybrid molecules. In the presence of ssRNA, a step of reverse transcription of the RNA molecule can be performed prior to PCR amplification.

One embodiment of the present invention is to determine the genetic profile of an individual or a patient comprising determining the presence of at least one polymorphic or haplotypic variation in UGT genes. The *UGT1A1*, *UGT1A7* and *UGT1A9* genes are the preferred candidate genes according to the present invention, where haplotypic variations can be found in a specific gene or considered simultaneously on multiple genes.

The putative UGT1A9 variations which can be investigated to determine a predisposition to a physiological reaction are $C^{-2208}T$ substitution, $C^{-2152}T$ substitution, $C^{-2141}T$ substitution, $T^{-1887}G$ substitution, $T^{-1818}C$ substitution, $C^{-665}T$ substitution, $T^{-440}C$ substitution, $C^{-331}T$ substitution, $T^{-275}A$ substitution, $G^{-87}A$ substitution, $G^{8}A$ missence mutation ($C^{3}Y$), a $T^{98}C$ missence mutation ($M^{33}T$), or a combination of these variations. The $G^{8}A$ missence mutation is generally associated with a decreased predisposition or susceptibility to an anti-cancer agent whereas the $T^{98}C$ missence mutation is associated with an increased predisposition or susceptibility to the same anti-cancer agent. Mutations that can be determined in the *UGT1A7* gene are $G^{353}T$ missense mutation ($G^{115}S$), $T^{397}G$ missense mutation ($N^{129}K$), $C^{401}A$ and $G^{402}A$ missense mutations ($R^{131}K$), $G^{427}C$ missense mutation ($E^{139}D$) or $T^{632}C$ missense mutation ($W^{208}R$), while the UGT1A1 variation is a TA_7 mutation in the TATA box. A person skilled in the art will recognize that any polymorphic or haplotypic variation found in a UGT gene that modify the expression of the UGT protein, its stability, its substrate specificity, its glucuronidation kinetic parameters or its primary, secondary, tertiary or quaternary structures also represents an aspect of the present invention.

The analysis of a nucleic acid molecule to identify a polymorphic or haplotypic variation can be performed by Restriction Fragment Length Polymorphism (RFLP) analysis, Allele Specific Oligonucleotide (ASO) analysis, Allele Specific PCR (ASP) analysis, Single Strand Conformation Polymorphism (SSCP) analysis, electronic microchip assay, denaturing high-performance liquid chromatography (DHPLC), allelic discrimination assays (Taqman), sequencing or using a DNA chip-based genotyping method, among others.

In one embodiment of the present invention, the analysis for determining a predisposition or a susceptibility to a drug, as for example but not limited to, an anti-cancer agent in a patient may be restrained to the analysis of UGT1A9 polymorphisms or combined with the analysis of other genes susceptible to lead to a predisposition or susceptibility to the anti-cancer agent (haplotype analysis). The latter genes may encode other drug-conjugating enzymes, such as UGT enzymes as described hereinabove, enzymes that mediate the

bioconversion of the CPT-11 molecule into SN-38 (carboxyesterase) or transport enzyme.

Since UGT1A1, UGT1A6, UGT1A7, UGT1A8 and UGT1A10 are the other UGT enzymes that conjugate CPT-11 and SN-38 molecules, the genes that encode these enzymes are targets used to investigate the glucuronidation haplotype of an individual, where at least one of these genes is analyzed concomitantly to UGT1A9. Polymorphic variations in other conjugating enzymes, belonging to the class of carboxyltransferases, sulfotransferases, glutathione S-transferase, methyltransferases or arylamine N-acetyltransferases, β-glucuronidases could also be investigated in concomitance to the UGT1A9 gene.

The transport enzymes described herein include, but are not limited to, ATP-binding cassette (ABC) proteins ABCA1, ABCA2, ABCA3, ABCA4, ABCA5, ABCA6, ABCA7, ABCA8, ABCA9, ABCA10, ABCA11, ABCA12, ABCA13, ABCA14, ABCB1, ABCB2, ABCB3, ABCB4, ABCB5, ABCB6, ABCB7, ABCB8, ABCB9, ABCB10, ABCB1, ABCC1, ABCC2, ABCC3, ABCC4, ABCC5, ABCC6, ABCC7, ABCC8, ABCC9, ABCC10, ABCC11, ABCC12, ABCC13, ABCD1, ABCD2, ABCD3, ABCD4, ABCE1, ABCF1, ABCF2, ABCF3, ABCG1, ABCG2, ABCG4, ABCG5, ABCG8, Breast cancer resistance protein (BCRP), multi-drug resistance protein (MRP) and PGY proteins.

As DNA repair mechanisms could be implicated in the hypersensitivity to camptothecin analogs, haplotype analysis that investigate these mechanism concomitantly to UGT haplotyping analysis is also one embodiment of the present invention. Genes that encode for DNA mismatch repair (MMR), homologous recombination (HR), non-homologous end joining (NHEJ) and single-strand annealing (SSA) systems, as well as Rad and ATPase proteins could therefore be analyzed by a skilled artisan simultaneously to UGT sequences.

In a further embodiment of the present invention, there is provided an isolated nucleotide molecule comprising an allelic variant of a polymorphic region of a

UGT1A1 gene, wherein the allelic variant comprises at least one TATA box polymorphic variation within the UGT1A1 promoter region.

According to another embodiment of the present invention, there is provided an isolated nucleotide molecule comprising an allelic variant of a polymorphic region of a UGT1A7 gene, wherein the allelic variant comprises at least one nucleotide sequence selected from the group consisting of those set forth in SEQ ID No: 60 to SEQ ID NO: 68, or the complement thereof.

Also, there is provided an isolated nucleotide molecule comprising an allelic variant of a polymorphic region of a UGT1A9 gene, wherein the allelic variant comprises at least one nucleotide sequence selected from the group consisting of those set forth in SEQ ID NO: 36 to SEQ ID NO: 59, or the complement thereof.

In a further embodiment, there is provided an isolated amino acid sequence comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO: 69, SEQ ID NO: 70, SEQ ID NO: 71 or a fragment thereof. These amino acid sequences may be encoded by a nucleotide sequence comprising at least one sequence selected from the group consisting of SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, a fragment or the complementary sequences thereof. Alternatively, the expression of the amino acid sequence may be regulated by a nucleotide sequence comprising at least one sequence selected from the group consisting of SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45 SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, a fragment or the complementary sequences thereof.

The present invention will be more readily understood by referring to the following examples which are given to illustrate the invention rather than to limit its scope.

EXAMPLE I

Distribution of SN-38-Glucuronide Formation in Human Liver Microsome Samples

To obtain statistical data on interindividual variation of SN-38 glucuronidation, we measured the SN-38-G formation, as currently known in the background art, with microsomes preparations from each patient liver sample. The glucuronide formation rates were regrouped into ranges and every sample was categorized within these ranges.

The following results show a mean for glucuronidation rate of 0.61 pmol/mg of protein/minute (Table 1). Data also indicate a substantive distribution of the glucuronidation rates. Fig. 4 illustrates the distribution of the glucuronidation rates obtained with liver samples.

TABLE 1
Statistical data of SN-38 glucuronide formation by human liver samples

Quantiles			Moments	
100.0%	maximum	1.9735	Mean	0.6117859
99.5%		1.9735	Std Dev	0.5014612
97.5%		1.9309	Std Err Mean	0.0723797
90.0%		1.5699	upper 95% Mean	0.7573947
75.0%	quartile	0.9279	lower 95% Mean	0.4661772
50.0%	median	0.4204	N	48
25.0%	quartile	0.2763	Sum Wgts	48
10.0%		0.1641	Sum	29.365725
2.5%		0.1113	Variance	0.2514633
0.5%		0.1063	Skewness	1.3394982
0.0%	minimum	0.1063	Kurtosis	0.6950696
			CV	81.966773

EXAMPLE II

Identification of UGT1A9 Variants

MATERIAL AND METHODS

DNA samples

DNA samples of 201 Caucasian subjects were obtained from the Quebec Family Study (QFS) (Simonen *et al.*, 2002, *Med. Sci. Sports Exerc.* 34: 1137-1142). Unrelated Caucasian subjects were recruited at the Massachusetts General Hospital (n=100) and genomic DNA from African-American subjects were kindly provided by Robert Millikan (Lineberger Comprehensive Cancer Center, School of Medicine, University of North Carolina, Chapel Hill, NC 27599-7435, USA) (n=20). These samples had been anonymized prior to their reception in our laboratory. All subjects have provided written consent for the use of their DNA for experimental purposes, and the present study was reviewed and approved by Institutional Review Boards (CHUL Research Center and Laval University).

Resequencing of the *UGT1A9* gene and genotyping

Polymerase chain reaction (PCR) was used to amplify the first exon of the *UGT1A9* gene. Three pairs of primers were designed to amplify overlapping fragments covering the first exons, a small portion of the 5'-flanking region and the intron-exon boundary (listed in Table 1). PCR amplification and DNA sequencing were performed according to protocols of Faucher *et al* (Faucher *et al.*, 2002, *Hum. Mol. Genet.* 11: 2077-2090). Amplicons were sequenced with an ABI 3700™ automated sequencer using Big Dye™ (Perkin Elmer™) dye primer chemistry. Samples were sequenced on both strands with nested primers listed in Table 2. Samples with ambiguous sequencing chromatograms and samples with single nucleotide polymorphisms (SNPs) were subjected to a second, independent amplification, followed by DNA sequencing. Sequences were analyzed with Staden preGap4 and Gap4 programs. These programs align sequence chromatograms and identify areas in which polymorphisms might be present. Each chromatogram was then evaluated individually to confirm variation in the sequences.

To determine the prevalence of UGT1A9 alleles in the population, a portion of the first exon, which includes the newly discovered polymorphisms, was amplified by PCR using specific oligonucleotides #37 and #38 (SEQ ID NO: 1 and 2). PCR amplifications were performed in a final reaction volume of 50 µL containing 25 ng of genomic DNA, 20 pmol of each primer, 1X reaction buffer, 100 µM dNTPs, 4 % DMSO and 2 U of the *Taq* DNA polymerase. The amplification conditions were: denaturation at 96°C for 5 min, 35 cycles of 30 sec at 94°C, 40 sec at 58°C and 1 min at 72°C, with a final extension step of 7 min at 72°C. Reactions were performed in a Perkin Elmer™ model 9700 thermal cycle. ASOs were designed to detect by hybridization the missense mutations in the UGT1A9 amplification products. Four ASOs were designed to specifically hybridize to the sequence corresponding to a G or an A at codon 3 (Fig. 5e) and a T or a C at codon 33 (Fig. 5f) and hybridization performed as previously described (Guillemette *et al.*, 2000, *Pharmacogenetics* 10: 629-644).

TABLE 2

Primer sequences for UGT1A9

Primers	Sequences	SEQ ID NO:
PCR amplification UGT1A9		
#37	5' - gtgctggtatttctccc	1
#38	5' - gtcaaaaatgtcattgtatgaacc	2
#39	5' - gatctggaccgggagttcaa	3
#40	5' - gtgtggctgtagagatcatact	4
#41	5' - catgcacttggaggaacatttatta	5
#42	5' - gagtacacgcattggcac	6
Direct sequencing UGT1A9		
#7	5' - ctcccacctactgtatc	11
#8	5' - gttcaaggcttttgccc	12
#9	5' - catttattatgccaccg	13
Allelic specific oligos UGT1A9		
C^3	5' - atggcttgcacagggt	14
Y^3	5' - atggcttacacagggt	15
M^{33}	5' - agtgcccatggatggga	16
T^{33}	5' - agtgcccacggatggga	17
Site-directed mutagenesis UGT1A9		
C^3 to Y^3 (Forward)	5' - gttctctgatggcttacacagggtggaccag	28
C^3 to Y^3 (Reverse)	5' - ctggtccaccctgtgtaagccatcagagaac	29
M^{33} to T^{33} (Forward)	5' - gctactggtagtgcccacggatgggagccactgg	30
M^{33} to T^{33} (Reverse)	5' - ccagtggctcccatccgtgggcactaccagtagc	31

Bold : nucleic acid polymorphism

Methods for UGT1A9 SNPs detection

UGT1A9 first exon was amplified in unrelated subjects. Allelic discrimination PCR was used to genotype UGT1A9 codons 3 and 33. The probe marked with FAM fluorochrome was designed to detect the *wild type* allele. The other probe used to detect the polymorphic alleles were marked with TET fluorochrome.

Duplicate filters were hybridized separately with the corresponding γ-^{32}P labeled oligonucleotides. The positive signals detected with both ASOs indicated heterozygous individuals for the polymorphism in contrast with a positive signal with one probe only, which indicated that the subject was homozygous.

Identification of missense mutations in the human *UGT1A9* first exon by direct sequencing of PCR products.

Specific primers were used to amplify the exon 1 *UGT1A9* (SEQ ID NO: 1 to SEQ ID NO: 6). The nonsynonymous polymorphisms in the third codon (C^3Y) (a) and in codon 33 ($M^{33}T$) of UGT1A9 (b) are illustrated in Fig. 6.

Functional analysis of the conjugating activity of UGTA9 variants

Microsomal fractions from HEK-293 cells stably expressing human UGT1A9*1, UGT1A9*2 and UGT1A9*3 were used in enzymatic assays. Reactions (100 µl volume) contained 50 mM Tris-HCl, pH 7.3, 10 mM $MgCl_2$, 100µg/mL phosphatidylcholine, 1 mM UDP-glucuronic acid , 40 to 60 µg of membrane protein. SN-38, MPA or other substrates were added in concentrations ranging from 1 to 200 µM and the reaction was incubated 30 min. at 37 °C with agitation. Human liver microsome were incubated in the same condition for control. Reaction was stopped by the addition of 200 µL MeOH + 1 % HCl 2N, followed by centrifugation at 14 000 rpm for 10 minutes. Supernatant was filtered through a 0.22 µm filter and 100 µL of water was added. For SN-38 and SN-38-glucuronide detection, 10 µL samples were injected on a liquid chromatographic system coupled to a fluorescence detector. Time-course experiments were realized to determine the linearity of the glucuronidation reaction. For determination of V_{max} and K_m, HEK-293 cells stably expressing UGT enzymes were incubated in the presence of varying SN-38 concentrations from 1 to 200 µM for the corresponding period of 30 min. All reactions rates were shown to be linear for these times.

A liquid chromatographic method was developed to quantify SN-38 glucuronidation of UGT cell line-derived microsomes and human liver microsomes. Samples were analyzed using high performance liquid chromatography (Alliance 2695, Waters, Milford, MA). Chromatographic separation was achieved with a Colombus C18 column 5-µm packing material, 50 x 3.2 mm (Phenomenex, Torrance, CA) using a two-solvent gradient system : A (water + 1 mM ammonium formate); B (MeOH + 1 mM ammonium formate). At a constant flow rate (0.7 ml/min), a linear gradient from 20 to 65 % B was run over 3 min, held 0.8 min and a second gradient until 95 % of B was run over 2 min and then re-equilibrated to 20 % B over 2 min. The effluent from the HPLC

system (Alliance 2695) was connected directly to a fluorescence detector (Waters, Milford, MA) using an excitation wavelength of 460 nm and an emission of 460 nm. Retention time for SN-38 and SN-38-glucuronide were respectively 3.5 and 4.6 min. Determination of the glucuronidation rates obtained with other substrates was performed as currently known in the art.

RESULTS

Identification of two novel missense mutations in the human *UGT1A9* gene and their distribution in healthy individuals.

The strategy used to identify polymorphisms in the *UGT1A9* gene was a PCR amplification of the exon 1, followed by direct DNA sequencing. Inclusion of a portion of the adjacent intron and 5'-flanking region in the PCR fragment was performed in order to assure the specific amplification of the *UGT1A9* gene. The UGT1A9 was resequenced on both strands for 35 subjects. DNA samples from Caucasian-American subjects was shown to contain one SNP, whereas an additional SNP was observed in an African-American subject. No insertion-deletion events were observed within the area sequenced.

The nucleotide change producing the first cSNP (SNPs in the coding region) was a change of a G to an A at nucleotide 8. The polymorphic change results in the substitution of Cysteine by a Tyrosine (C^3Y) in the signal peptide of the UGT1A9 protein corresponding to the UGT1A9*2 allele (SEQ ID NO: 37). The second nucleotide change, $T^{98}C$, leads to a Methionine to a Threonine at codon 33 ($M^{33}T$) corresponding to the UGT1A9*3 allele (SEQ ID NO: 38) . Figs. 6a and 6b illustrate the sequence analysis of three genotypes: homozygous *wild type* *1/*1 and heterozygous *1/*2 or *1/*3.

To determine the allelic frequency of UGT1A9 allozymes in the population, we genotyped unrelated subjects including 301 Caucasians of whom 201 were French-Canadians, and 20 African-American subjects. Only one African-American individual had the C^3Y mutation whereas 12 individuals, all Caucasian subjects, were shown to have the $M^{33}T$ mutation (illustrated in Fig. 5f). A total of 5 % of individuals were found heterozygous for the UGT1A9*3 allele in the French-Canadian population and 3 % of the remaining Caucasian-American

subjects. None of the 20 African-American subjects were found to have the UGT1A9*3 allele (Table 3).

TABLE 3
Allelic frequency and prevalence of UGT1A9 alleles

		Allele frequency			Genotype frequency (%)[a]		
	n	*1	*2	*3	*1/*1	*1/*2	*1/*3
Amino acid change		Cys^3Met^{33}	Tyr^3Met^{33}	Cys^3Thr^{33}			
Functional change		"Wild type"	Similar activity	Decreased activity			
Population characteristics							
Caucasian (French Canadian)	201	0.978	0.000	0.022	95	0	4.4
Caucasian (American)	100	0.964	0.000	0.036	97	0	3
African (American)	20	0.975	0.025	0.000	95	5	0

[a] Subjects homozygous for variant UGT1A9 alleles were not observed in the population tested.

Functional analysis of the conjugating activity of UGTA9 variants

Table 4 shows that the presence of a threonine at position 33 (UGT1A9*3) is correlated to 96.3% decreased conjugation rate for SN-38 while the presence of a tyrosine at codon 3 is associated to a 16.7% increased activity. Moreover, modulation of the UGT1A9 glucuronidation activity is substrate specific since conjugation of eugenol, 2-hydroxyestradiol, 4-hydroxyestrone and 4 methylumbelliferone is increased or decreased in a proper way for each substrate. The presence of a threonine at position 33 does not affect significantly the affinity of the protein for SN-38 but decreases by approximately 20-folds its glucuronidation rate (Table 5) while the affinity of UGT1A9 for MPA is dramatically reduced by the presence of codon 33 variation (Table 6).

TABLE 4

Substrate-dependent modulation of the UGT1A9 activity by codons 3 and 33.

Substrates	% glucuronide formation relative to UGT1A9*1	
	UGT1A9*2	UGT1A9*3

Eugenol	↓ 28.4 %	↑ 716 %
2-OH-E2	↓ 26.6 %	↑ 1727 %
4-OH-E1	↓ 19.0 %	↓ 90.3 %
4-MU	↓ 19.2 %	↓ 66.2 %
SN-38	↑ 16.7 %	↓ 96.3 %
Flavopiridol	ns	ns

TABLE 5

Kinetic analysis of SN-38 glucuronidation by UGT1A9*1, *2 and *3

	1A9*1	1A9*2	1A9*3
Km	3.03 ± 0.72	5.15 ± 1.81	3.21 ± 0.95
Vmax	316.34 ± 52.03	324.68 ± 95.09	15.50 ± 8.40 $p < 0.001$
Vmax / Km (Cl_{int})	104	63	5

TABLE 6

Kinetic analysis of MPA glucuronidation by UGT1A9*1, *2 and *3

UGT1A9 alleles	Km µM	Vmax pmol/min/mg	Vmax/Km
1A9*1	495	9406	19
1A9*2	303	8401	28
1A9*3	3225	14074	4

EXAMPLE III

Identification of novel UGT1A9 promoter variants

The primary objective of this study was to examine the genomic sequences of the *UGT1A9* gene promoter sequence to identify novel expression

polymorphisms and to determine whether or not these polymorphic variations would affect the expression of the UGT1A9 protein. To determine the effect of the polymorphic variations on the UGT1A9 protein expression, semi-quantitative immunoblot analyses were performed on liver microsomes from patients and correlated with their genotypes. Identification of novel polymorphisms has been performed by direct sequencing of a pool of DNA samples from patients. Determination of genotypes of each patient monitored was also performed by direct sequencing.

Liver microsomes from patients were prepared by differential centrifugation. The crude cell extracts were centrifuged at 12 000 x *g* at 4°C for 22 min to remove nuclei and other cellular debris. Supernatants were centrifuged at 105 000 x *g* for 60 min at 4°C to obtain the membrane fraction, which was homogenized in the buffer described above. Protein concentrations were determined using the Bradford method according to the manufacturer's recommendations.

To determine the level of UGT1A9 proteins expressed in the microsomal fractions obtained from liver microsomes, Western blot analyses were conducted as follows: Microsomal proteins (10 µg) from liver microsomes were separated by 10 % SDS-polyacrylamide gel electrophoresis. The separated proteins were transferred onto nitrocellulose membranes and probed with the antihuman UGT1A antiserum (1:1000 dilution) specific for the amino-terminal region of the UGT1A7, UGT1A8, UGT1A9 and UGT1A10 proteins. Given that UGT1A7, UGT1A8 and UGT1A10 are not expressed in liver tissue, immunodetection with this antiserum in human liver microsomes is specific to UGT1A9. In order to normalize sample loading, blots were re-probed with anti-calnexin antibody (1:2000 dilution; StressGen Biotechnologies Corp., Victoria, Canada), to detect a second ER-resident protein. A donkey antirabbit IgG antibody conjugated with the horseradish peroxidase (Amersham Corp., Oakville, Canada) was used as the secondary antibody (1:10 000 dilution). The resulting immunocomplexes were visualized using an enhanced chemiluminescence kit (ECL) (Renaissance, Quebec, Canada) and exposed on Kodak XB-1 film. The lowest signal has been used as standard to determine

the relative expression of UGT1A9 in each sample and results were monitored by Oneway analyses.

RESULTS

Ten novel polymorphic variations were identified within the *UGT1A9* promoter region, namely a C(-2208)T substitution, a C(-2152)T substitution, a C(-2141)T substitution, a T(-1887)G substitution, a T(-1818)C substitution, a C(-665)T substitution, a T(-440)C substitution, a C(-331)T substitution, a T(-275)A substitution and a, G(-87)A substitution.

UGT1A9 protein expression is highly variable among tested samples, as shown on Fig. 7. Figs. 8a to 8e demonstrate a positive correlations between the presence of mutated nucleic acids in positions -2152 (Fig. 8a), -665 (Fig. 8b), -440 (Fig. 8c), -331 (Fig. 8d) and -275 (Fig. 8e) in the promoter region of the *UGT1A9* gene and the expression of higher level of UGT1A9 proteins.

EXAMPLE IV

Effect of *UGT1A9* polymorphic variations on liver microsomes glucuronidation

One it has been established that polymorphic variations in the promoter region of the UGT1A9 gene can modulated the expression of the UGT1A9, it was interesting to study the impact of these mutation on global glucuronidation by human liver microsomes. Therefore, a correlation study was undergone to determine if correlations could exist between C(-2152)T, T(-1818)C, C(-665)T and T(-275)A variations and SN-38, mycophenolic acid and 4-hydroxyestrone glucuronide formation. Glucuronidation activity was determined for each liver sample in nmoles/mg of proteins/min and further regrouped respective to the genotype of the patient, namely patient carrying a mutation or non-carrying (wild type) patients.

Results

One way analyses demonstrate a correlative association between the presence of a mutated nucleic acid at position -2152 and glucuronidation of MPA (Fig. 9). Fig. 10 also shows a positive correlation between the formation of SN-38-glucuronide and the presence of one or both mutated alleles at position -1818 in the UGT1A9 promoter region. Nucleic acid change at position -665 correlates with higher glucuronidation rates with SN-38, (Fig. 11a), 4-hydroxyestrone (Fig. 11b) and mycophenolic acid (Fig. 11c). Finally, Fig. 12 shows a positive correlation between the presence of the -275 mutated alleles and higher glucuronidation rate with SN-38.

EXAMPLE V

Effect of the expression of UGT1A proteins on glucuronidation by liver microsomes

As UGTA9 is considered as a major SN-38 glucuronidation enzyme, we attempted to determine if an association between the expression of this proteins and glucuronide formation could exist. As shown in Fig. 13a, there is a positive correlation between glucuronidation of SN-38 and protein level of UGT1A9. To ascertain that the enhancement of glucuronidation observed with this substrate is not attributable to a residual activity of other UGT isoforms, these experiments were reconducted using a probe substrate for UGT1A9, namely mycophenolic acid. Fig. 13b illustrates the positive correlation between UGT1A9 protein expression level and MPA glucuronidation.

EXAMPLE VI

Identification of novel UGT1A7 variants

The primary objective of the study was to examine the genomic sequences of the *UGT1A7* gene, for which functional polymorphisms have been described yet to identify novel polymorphic variations. The aim was to look for missense polymorphisms in a Caucasian population, to develop methods for SNPs detection and to evaluate their functional properties after *in vitro* expression of

enzyme variants. In turn, *UGT1A7* is a polymorphic gene for which there are at present four known allelic variants (Guillemette *et al.*, 2000, *Pharmacogenetics*, 10: 629-640). Based on *in vitro* metabolic studies, the UGT1A7*3 and *4 variants may potentially lead to a poor SN-38 glucuronidator phenotype.

MATERIAL AND METHODS

UGT1A7 haplotype determination

DNA samples were obtained according to Example 2. To discriminate the polymorphisms at codons 129/131 and 139, a PCR technique using the Taqman® technology was used (Applied Biosystems, Branchburg, NJ, USA). To discriminate the two alleles at codons 129/131, the exon 1 containing the codon 129/131 was amplified using primers 387 and 388 (SEQ ID NO: 20 and 21, respectively) shown in Table 4. Two probes were designed to identify the two different alleles, probe for N^{129}/R^{131} allele was marked with FAM fluorochrome and probe for K^{129}/K^{131} allele was marked with TET fluorochrome. Also, specific primers were designed to amplify the region of exon 1 containing codon 139. Specific 21-mer probes were designed to identify the two different alleles. One of the probes, E^{139}-FAM, was homologous to the *wild type* allele. The other probe, D^{139}-VIC, contained the polymorphic nucleotide at codon 139 in order to be homologous to the D^{139} mutant allele. Each PCR reaction was performed with 25 ng of genomic DNA in a volume of 10 µL and containing 5 pmole of each primer and probe and 1 x Taqman® universal PCR master mix. PCR conditions were 50°C for 2 minutes, 95°C for 10 minutes followed by 40 cycles at 95°C for 15 seconds and 60°C for 1 minute. The ABI prism 7000™ system detected the different genotypes (Figs. 5a; 5c).

The polymorphism at codon 208 of UGT1A7 was genotyped by PCR-RFLP. The polymorphism at codon 208 creates a restriction site for *Rsa* I enzyme. Digestion was performed with 5 µL of PCR product, 10 U of *Rsa* I and 1 x reaction buffer L (10mM Tris-HCl, 10mM $MgCl_2$, 1mM DTE, PH 7.5) in a total volume of 10 µL. Reactions were incubated for 2 hours at 37°C and separated on a 2% agarose gel to observe the different migration patterns. Homozygous *wild type* genotype at codon 208 generates a single fragment migrating at 590 bp. The heterozygous genotype generates a fragment of 590 pb representing

the *wild type* allele and two bands of 236 and 264 bp representing the polymorphic allele cut by *Rsa* I. Homozygous mutants at position 208 have a pattern of migration showing only two bands of 236 and 264 bp (Fig. 5b).

Allelic specific oligonucleotides (ASOs) were designed to detect UGT1A7 polymorphism at codon 115. PCR amplification using primers 292 and 293 was used to generate the target fragment containing the polymorphic site. Each ASO is composed of a 17-mer centered over the polymorphic nucleotide of each variant. The denatured PCR products were spotted onto filters, each one being subsequently hybridized with a single ASO using a method that has been described previously (Guillemette *et al.*, 2000, *Pharmacogenetics*, 10: 629-640). Conditions for ASO hybridization analysis have been described above and a typical result is illustrated in Fig. 5d.

Methods for UGT1A7 SNPs detection.

UGT1A7 first exon was amplified in unrelated subjects. (a) Allelic discrimination PCR was used to genotype UGT1A7 codons 129/131. The probe marked with FAM fluorochrome was designed to detect the *wild type* N^{129}/R^{131} allele. The other probe used to detect the polymorphic allele K^{129}/K^{131} was marked with TET fluorochrome. (b) PCR products amplified with primers #17 (SEQ ID No: 8) and #18 (SEQ ID No: 7) were digested using *Rsa* I enzyme to determine whether the patients were homozygous *wild type* W^{208}, heterozygous W^{208}/R^{208} or homozygous R^{208}. The 590 bp fragment represents the undigested PCR product whereas the 336 and 264 bp fragments result from the digestion of the 590 bp amplicon. (c) Allelic discrimination PCR was used to genotype the novel polymorphism at codon 139 of the *UGT1A7* gene. The FAM fluorochrome was used to mark the *wild type* probe E^{139} and the VIC fluorochrome was used for the polymorphic probe D^{139}. (d) Allelic specific oligonucleotides (ASOs) were designed to genotype the novel polymorphic variation at codon 115 of *UGT1A7* gene. (e) (f) A similar strategy was further used to detect variants at codons 3 and 33 of the *UGT1A9* gene. Duplicate filters were hybridized separately with the corresponding γ-^{32}P labeled oligonucleotides. The positive signals detected with both ASOs indicated heterozygous individuals for the polymorphism in

contrast with a positive signal with one probe only, which indicated that the subject was homozygous.

Identification of missense mutations in the human *UGT1A7* first exon by direct sequencing of PCR products.

Specific primers were used to amplify the exon 1 of *UGT1A7* (Table 7). The nonsynonymous polymorphisms illustrated, along with the codon 115 ($G^{115}S$) (c) and codon 139 ($E^{139}D$) (d) polymorphisms of UGT1A7 (see Fig. 6). The sequence illustrated in (a) and (b) correspond to the "sense" strand whereas (c) and (d) correspond to the "anti-sense" strand.

TABLE 7

Primer sequences for UGT1A7

Primers	Sequences	SEQ ID NO:
PCR amplification UGT1A7		
#18	5'- cgctggacggcaccattg	7
#17	5'- gctaaaggggagataacttacc	8
#122	5'- gctggacggcaccattg	9
#123	5'- ccctaagagaagtctgggg	10
Allelic specific oligos UGT1A7		
G^{115}	5'- catccaatggtattttt	18
S^{115}	5' -catccaatagtattttt	19
Taqman® analysis (Codon 129/131) UGT1A7		
#387	5'- gcaccattgcgaagtgcat	20
#388	5'- ggatcgagaaacactgcatcaa	21
N129/R131-FAM	5'- ttaatgaccgaaaatt	22
K129/K131-TET	5'- tttaaggacaaaaaatt	23
Taqman® analysis (Codon 139) UGT1A7		
#546	5'- gcgaagtgcattttctctattaacaa	24
#544	5'- aagccacagcgatcaaaagg	25
E139-Fam	5'- atacttaaaggagagttgttt	26
D139-Vic	5'- atacttaaaggacagttgttt	27
Site-directed mutagenesis UGT1A7		
E^{139} to D^{139}(Forward)	5'- aattagtagaatacttaaaggacagttgttttgatgcagtgtttc	32
E^{139} to D^{139}(Reverse)	5'- gaaacactgcatcaaaacaactgtcctttaagtattctactaatt	33
G^{115} to S^{115}(Forward)	5'- gttcatccaatagtatttttgac	34
G^{115} to S^{115}(Reverse)	5'- gtcaaaaatactattggatgaac	35

Bold : nucleic acid polymorphism

RESULTS

Identification of two novel polymorphisms in the coding region of the *UGT1A7* gene and haplotypic structure analysis of the *UGT1A7* gene.

The exon 1 of UGT1A7 was amplified by PCR in 117 subjects, 54 Caucasians and 63 African-Americans, and then sequenced. Two novel polymorphisms were found at codon 115 and 139 (Figs. 6c; 6d). At codon 115, a nucleotide change of a G to an A leads to an amino acid change from Glycine to Serine ($G^{115}S$). A G to C mutation at codon 139 leads to an amino acid change from Glutamate to Aspartate ($E^{139}D$).

When combined with the previously described variations at codon 129/131 and 208, nine haplotypes were found to exist (UGT1A7 *1 to *9, Table 4). Four alleles were previously described, *1 to *4, and novel alleles correspond to UGT1A7*5 $S^{115}N^{129}R^{131}E^{139}W^{208}$ (Genbank AF434903), UGT1A7*6 $G^{115}N^{129}R^{131}D^{139}W^{208}$ (Genbank AF434904), UGT1A7*7 $G^{115}K^{129}K^{131}D^{139}W^{208}$ (Genbank AF461758), UGT1A7*8 $G^{115}K^{129}K^{131}D^{139}R^{208}$ (Genbank AF436810) and UGT1A7*9 $S^{115}K^{129}K^{131}E^{139}W^{208}$ (Genbank AF463483).

According to their prevalence in the population tested, the nine variant alleles were separated in two categories: the common and the rare alleles. The common alleles *1, *2 and *3, are present at a allelic frequency of 0.31 to 0.32. The rare alleles are UGT1A7*4 to *9, with frequencies between 0.002 to 0.025. The allelic frequencies for the polymorphisms at codon 115 and 139 were 0.04 and 0.06, respectively and found specifically in African-American individuals.

TABLE 8

Allelic frequency and prevalence of UGT1A7 alleles

UGT1A7 alleles		Function[b]	Frequency[c]
UGT1A7*1[a]		High	32.18
UGT1A7*2	K^{129}/K^{131}	High	30.60
UGT1A7*3	$K^{129}/K^{131}/R^{208}$	Low	31.55
UGT1A7*4	R^{208}	Low	2.52
UGT1A7*5	S^{115}	Low	0.47
UGT1A7*6	D^{139}	High	0.16
UGT1A7*7	$K^{129}/K^{131}/D^{139}$	High	2.06
UGT1A7*8	$K^{129}/K^{131}/D^{139}/R^{208}$	Low	0.16
UGT1A7*9	$S^{115}/K^{129}/K^{131}$	Low	0.32

[a] UGT1A7*1: $G^{115}/N^{129}/R^{131}/E^{139}/W^{208}$; only position differing from *1 are indicated

[b] Based on in vitro experiments: Low : significantly lower SN-38G formation versus *1 allele.
High : no significant difference in activity compared to *1 allele.

[c] Population of 167 Caucasian and 150 African-American subjects.

TABLE 9
Frequency of the UGT1A7 alleles

UGT1A7 genotypes[a]	Population (n=317)[b]	Frequency (%)
*1/*1	30	9.46
/*2	57	17.98
*1/*3 *2/*4	72	22.71
*1/*4	6	1.89
*1/*6	1	0.32
*1/*7 *2/*6	7	2.21
*1/*8 *3/*6 *4/*7	1	0.32
*2/*2	39	12.30
*2/*3	55	17.35
*2/*7	3	0.95
*2/*8 *3/*7	1	0.32
*2/*9	1	0.32
*3/*3	35	11:04
*3/*7	2	0.63
*4/*4	5	1.58
*5/*5	1	0.32
*5/*9	1	0.32
Low activity genotypes[c]	42	13.26
Intermediate activity genotypes[d]	138	43.54

[a] In bold: Genotypes considered to evaluate allelic frequencies.
[b] 167/317 Caucasian ; 150/317 African-American subjects.
c With two low activity alleles.
[d] With one low activity allele.

EXAMPLE IV

Relative expression of the UGT1A7 and UGT1A9 variants and SN-38 glucuronidation activities of UGT1A7 and UGT1A9 allozymes

MATERIAL AND METHODS

UGT1A7 and UGT1A9 expression studies

All five novel UGT1A7 variant alleles were generated by PCR site-directed mutagenesis using pcDNA3-vector containing either UGT1A7*1, *2, *3 or *4 variant alleles as the starting construction. Primers having SEQ ID NO: 32, 33, 34 and 35 (Table 7) were used for site-directed mutagenesis. The variant

alleles *5 (SEQ ID NO: 50) and *6 (SEQ ID NO: 51) were generated using *1 (SEQ ID NO: 46) as the template, the *7 (SEQ ID NO: 52) and *9 (SEQ ID NO: 54) variants were obtained using the *2 (SEQ ID NO: 47) allele as template and the *8 (SEQ ID NO: 53) was created from *3 (SEQ ID NO: 48) allele. Expression constructs for the UGT1A9 cDNA sequence construct and constructs for the two nonsynonymous cSNPs were created using the same strategy. The expression plasmid pcDNA3-UGT1A9*1 was obtained by subcloning the *Bam* HI-*Xho* I fragment of pBK-CMV / UGT1A9*1 (kindly provided by Dr Alain Belanger from CHUL Research Center, Laval University, Québec, Canada) into the *Bam* HI-*Xho* I site of pcDNA3 expression vector. Mutations were all verified by sequencing. Stable HEK293 cells were transfected with variant pcDNA3-UGT1A7 and pcDNA3-UGT1A9 expression plasmids using the following procedure that has been described previously (Guillemette *et al.*, 2000, *Pharmacogenetics*, 10: 629-640). HEK293 cells in the exponential growth phase were seeded at a density of 3.25×10^6 cells/culture dish. Briefly, cells were grown in Dulbecco's-modified Eagle's medium (DMEM) containing 10 % fetal bovine serum (FBS), 1 % Sodium Pyruvate (NaPy) and 0.1 mg/mL Amikacin in a humidified incubator at 37°C with an atmosphere of 5 % CO_2. The next day, cells at 60 % of confluence were washed with DMEM without FBS. Then, cells were incubated with 5 mL of the same medium containing 30 µL Exgen 500™ (MBI fermentas, Burlington, ON, Canada) and 15 µg of the appropriate pcDNA3-UGT expression plasmids. Transfections were stopped after 3 hours by the addition of fresh DMEM with 10 % FBS. After 48 hours, geneticin (1 mg/mL) (Invitrogen life technologies, Carlsbad, CA) was added to begin the selection process. During the following 4 weeks, fresh medium with antibiotic was added every 2 days until colonies of resistant cells became visible and for amplification of geneticin-resistant cell populations.

Microsomes were prepared by differential centrifugation. The crude cell extracts were centrifuged at 12 000 x *g* at 4°C for 22 min to remove nuclei and other cellular debris. Supernatants were centrifuged at 105 000 x *g* for 60 min at 4°C to obtain the membrane fraction, which was homogenized in the buffer

described above. Protein concentrations were determined using the Bradford method according to the manufacturer's recommendations.

To determine the level of UGT proteins expressed in the microsomal fractions obtained from the stably transfected cells, Western blot analyses were conducted as follows. Microsomal proteins (10 µg) from HEK293 cells stably expressing human UGT1A9 and UGT1A7 variants were separated by 10 % SDS-polyacrylamide gel electrophoresis. The separated proteins were transferred onto nitrocellulose membranes and probed with the antihuman UGT1A antiserum RC71 (1:1000 dilution) specific for the conserved C-terminal region of the protein. In order to normalize sample loading, blots were re-probed with anti-calnexin antibody (1:2000 dilution; StressGen Biotechnologies Corp., Victoria, Canada), to detect a second ER-resident protein. A donkey antirabbit IgG antibody conjugated with the horseradish peroxidase (Amersham Corp., Oakville, Canada) was used as the secondary antibody (1:10 000 dilution). The resulting immunocomplexes were visualized using an enhanced chemiluminescence kit (ECL) (Renaissance, Quebec, Canada) and exposed on Kodak™ XB-1 film. The relative levels of UGT1A allozymes and calnexin were determined by integrated optical density (IOD) using Bioimage programs visage 110S (Genomic solution inc., Ann Arbor, MI, USA) and compared to the *1 respective UGT1A9 (SEQ ID NO: 36) and UGT1A7 (SEQ ID NO: 60) alleles.

Western blot analyses of UGT1A7 and UGT1A9 variants expressed in HEK293 cells were performed on microsomal proteins (10 µg) separated on a 10 % SDS-polyacrylamide gel. After transferring the proteins, the membranes were probed with an anti-UGT1A RC-71 polyclonal antibody and with an anti-calnexin antibody. The relative levels of UGT1A9 (a) and UGT1A7 proteins (b) were determined by semi-quantitative densitometric analysis of the Enhanced chemiluminescence (ECL) image. The *in vitro* SN-38 activity was assessed using microsomal fractions prepared from HEK293 cells expressing the *1 and variant UGT1A9 (c) and UGT1A7 (d) alleles and incubated with 5 µM of SN-38 as described in Materials and Methods.

RESULTS

Recombinant allozyme Western blot analysis.

Semi-quantitative Western blot analyses (Figs. 14a and 14b) showed high levels of immunoreactive UGT protein in all membrane fractions from HEK293 cell lines stably expressing UGTs. An anti-calnexin polyclonal antibody was also used in combination as an internal reference. Significant expression of all UGT1A7 and UGT1A9 alleles was found adequate allowing enzymatic assays to be performed.

EXAMPLE IV

Loss of function variants of the UGT1A7 and UGT1A9 enzymes

MATERIAL AND METHODS

Enzyme assays

Recombinant allozymes were assayed for UGT activity with the two anticancer agents, SN-38 and flavopiridol, as substrates. Microsomal fractions from HEK293 (40 to 60 µg) were added to a reaction mixture (100 µL) containing 50 mM Tris-HCl, pH 7.3, 10 mM $MgCl_2$, 100 µg/mL phosphatidylcholine and 2 mM UDP-glucuronic acid. SN-38 was added in concentrations ranging from 0.1 to 200 µM whereas flavopiridol was used at two concentrations: 5 and 200 µM. Commercially available human liver microsomes (Human Cell Culture Center Inc., Laurel, MD) were incubated in the same conditions for all experiments. Time-course experiments were performed to determine the linearity of the glucuronidation reaction. For the determination of V_{max} and K_m, HEK293 cells stably expressing UGT1A9 enzymes were incubated in the presence of various concentrations of SN-38 ranging from 0.1 to 200 µM and incubated for 30 min as described above whereas UGT1A7 membranes preparations were incubated for 3 hours. All reaction rates were shown to be linear in these conditions. Reactions with SN-38 were stopped by the addition of 200 µL MeOH + 1 % HCl 2N, followed by centrifugation at 14 000 x *g* for 10 minutes. The supernatants were filtered through a 0.22 µm membrane and 100 µL of water was added to the filtrate. For the detection of SN-38 and its glucuronide (SN-38G), 10 µL

samples were injected in a liquid chromatographic system (HPLC) coupled to fluorescence detector as described below.

A HPLC method was developed to quantify the rates of SN-38 glucuronidation from the various microsomal fractions under study. The HPLC system used was an Alliance 2695 (Waters, Milford, MA) equipped with a 50 x 3.2 mm Colombus C18 column (Phenomenex, Torrance, CA). The chromatographic separation was achieved with a two-solvent gradient system: solvent A (water + 1 mM ammonium formate); solvent B (MeOH + 1 mM ammonium formate). A linear gradient starting at 20 % solvent B was generated over a 3 min period and at a constant flow rate (0.7 mL/min) until a plateau was reached at 65% solvent B and held for 0.8 min. Then a second gradient ranging from 65% to 95% solvent B was generated during the following 2 min. Finally, the column was re-equilibrated to 20 % solvent B for 2 min. The column was connected to a fluorescence detector model 474 (Waters, Milford, MA) and the molecules were excited at a wavelength of 370 nm and an emission of 425 nm. The retention times for SN-38 and SN-38G were 4.49 and 3.12 min, respectively. Because we could not perform kinetic analysis with the UGT1A9*3 using the previously used electrospray ion-trap mass spectrometry method, the fluorescence detection was preferred since it was more sensitive in these conditions and allowed the detection of SN-38G formed by UGT1A9*3 microsomes at low concentrations of SN-38. K_m calculated for the human liver microsomes using both analytical methods were shown to be similar (6.8 ± 3.0 µM with the LCQ detector and 4.8 ± 0.8 µM with the fluorescent detection (data not shown)). Glucuronidation assay using flavopiridol as substrate were performed as previously described (Ramirez *et al.*, 2002, *Pharm. Res.* 19: 588-594). Relative glucuronidation activities for flavopiridol (5 and 7 glucuronides) were determined for one hour using 5 and 200 µM of substrate and in the same experimental conditions as used for SN-38.

RESULTS

Recombinant UGT1A7 and UGT1A9 enzyme SN-38 kinetics.

The functional genomic studies were focused on two anticancer drugs, SN-38 and Flavopiridol. UGT1A7 was previously shown to have the highest intrinsic clearance with SN-38 as substrate along with UGT1A1 and UGT1A9 (Gagne *et al.*, 2002, *Mol. Pharmacol.* 62:608-617) whereas UGT1A9 is the main UGT involved in the metabolism of flavopiridol (Ramirez *et al.*, 2002, *Pharm. Res.* 19: 588-594).

The chromatograms obtained after separation of the reaction products following enzymatic assays with 5 µM of SN-38 and the UGT1A9 variant allozyme preparations are depicted in Fig. 15 a), b) and c) *3. The formation of SN-38G by the UGT1A9*3 enzyme is markedly reduced, with only 3.8 % residual activity compared to the *wild type* enzyme (Fig. 15c). Our results thus demonstrate that the $M^{33}T$ polymorphism dramatically impairs the conjugation rate of SN-38 whereas no significant effect was observed with the UGT1A9*2 allozyme. In contrast, the formation of flavopiridol-G was not statistically different for UGT1A9*2 and UGT1A9*3 compared to the UGT1A9*1 allele at both low and high concentrations (5 µM and 200 µM of flavopiridol), suggesting a substrate specific impact of this amino acid variation in the UGT1A9 protein.

To determine if the amino acid change at codon 33 affects enzyme activity by an alteration of kinetic properties, glucuronidating activity of UGT1A9 allozymes was assessed using a wide range of SN-38 concentrations (0.1 to 200 µM). A non significant higher apparent K_m value for the UGT1A9*2 variants was observed as determined at least in three independent experiments. Both UGT1A9*1 and UGT1A9*3 alleles demonstrated a similar apparent K_m of 3.03 ± 0.51 and 3.21 ± 0.95, respectively (Table 9). As a result, decreases in level of enzyme activity observed for the UGT1A9*3 allele could not be attributed to the alterations of substrate affinity. However V_{max} values were about 26 times lower for UGT1A9*3 compared with UGT1A9*1 (11.89 ± 2.61 versus 316.34 ± 52.03 pmol/min/mg of protein, $p < 0.002$).

In the analysis of UGT1A7 allozymes, the highest SN-38 glucuronidating activity was observed for UGT1A7*1, *2, *6 and *9. Three novel low activity alleles were identified and the *5, *7 and *8 alleles presented 38-76% lower rates of SN-38G formation compared to UGT1A7*1, similar to the range of activity of the *3 and *4 alleles previously identified as low SN-38 glucuronidating activity alleles (Gagne *et al.*, 2002, *Mol. Pharmacol.* 62:608-617).

TABLE 10

Kinetic parameters for SN-38 glucuronidation by human UGT1A9 allozymes

UGT1A9 allozymes	Apparent K_m (μM)	V_{max} (pmol/min/mg protein)	Catalytic efficiencies V_{max}/K_m (μL/h/mg)
UGT1A9*1	3.02 ± 0.51	316.34 ± 52.03	105
UGT1A9*2	5.15 ± 1.81	324.38 ± 95.09	63
UGT1A9*3	3.21 ± 0.95	11.89 ± 2.61 *	4

The values of apparent K_m and V_{max} for the formation of SN-38 glucuronide were determined using microsomal preparations from UGT1A9-HEK293 cells. Values were expressed as the mean ± SD of at least three independent experiments performed in duplicate from Lineweaver-Burk plots. * $p < 0.002$ compared to UGT1A9*1.

EXAMPLE V

Immunofluorescence localization of UGT1A9*1, UGT1A9*2 and UGT1A9*3 proteins.

MATERIAL AND METHODS

Immunofluorescence visualization

One cSNP found in the UGT1A9 first exon was located in the signal peptide, thus immunofluorescence experiments were designed to localize the expressed protein within the cells. Stable HEK293 cells expressing human UGT1A9*1, UGT1A9*2 and UGT1A9*3 and also with cells transfected with pcDNA3 vector

alone were seeded on culture slides (VWR Scientific, West Chester, PA) and allowed to grow for 18 h. Then, cells were washed three times with PBS and fixed for 20 min with paraformaldehyde 2 % (w/v, Sigma, St. Louis, MO) in PBS. The slides were washed three times with PBS before permeabilization of the membranes for 40 min in PBS containing Saponin 0.2 % (w/v, Sigma, St. Louis, MO). After three washes with PBS, the cells were incubated for 30 min with gelatin 0.2 % in PBS (w/v, Sigma, St. Louis, MO). The permeabilized cells were incubated with a rabbit anti-UGT1A primary antibody (RC-71) at a 1:1000 dilution (v/v) in PBS containing Saponin 0.1 % and bovine serum albumin 1.5 %. Slides were incubated for 1 h and then washed three times with PBS. A goat anti-rabbit secondary antibody (Alexa Fluor 488, Molecular Probes Inc., Eugene, OR) was added at a 1:400 dilution in the same buffer as the primary antibody, and slides were incubated for 30 min at room temperature in the dark. Cells were then washed three times with PBS. Cell counterstaining was achieved by incubating the slides for 30 sec in the dark at room temperature with a 1:1000 (v/v) dilution of diamidino-2-phenylindole (DAPI, Molecular Probes Inc., Eugene, OR). Finally, cells were washed with PBS and mounted with a mounting medium (Sigma, St. Louis, MO). For visualization, a Fluoview confocal microscope (BX-61, Olympus, Melville, NY) with a 100 X oil objective was used.

On Fig. 16, HEK293 cells stably expressing pcDNA3 (a) or human UGT1A9 alleles (d), (g), (j) were fixed, permeabilized and then treated with a rabbit anti-UGT1A primary antibody (RC-71), followed by a goat anti-rabbit secondary antibody. Cell counterstaining of the nuclei was performed using DAPI (b), (e), (h), (k). To confirm the localization of the UGT proteins, a combination of the images obtained with the antibodies and the counterstain are shown in (c), (f), (i), (l).

RESULTS

To determine if the subcellular localization of UGT1A9 was affected by the codon 3 mutational polymorphism in the signal peptide region, immunofluorescence experiments were carried out. Coloration with diamidino-2-phenylindole (DAPI) was restricted to the nucleus (Figs. 16e, h and k) whereas the low background observed in the pcDNA3 control vector is due to

autofluorescence (Figs. 16a, b and c). UGT1A9*1, UGT1A9*2 and UGT1A9*3 proteins were localized in the cytoplasm and the perinuclear zone as well as in the endoplasmic reticulum (Figs. 16d, g and j).

EXAMPLE VI

Effect of UGT1A1 TATA box variations on UGT1A1 protein expression and glucuronidation activity

Although a correlative association between TATA box polymorphic variation is reported in prior art, the UGT1A7 and UGT1A9 interindividual variations of the present invention remained unknown at this time and their effect on SN-38 glucuronidation therefore remained unconsidered. In an attempt to decipher the particular function of every participating isoform in SN-38-G formation, we were interested to determine whether or not a correlative association could be made between the number of TA repeat in the TATA box of UGT1A1 promoter region and UGT1A1 protein expression even thought novel polymorphic variations were taken into account. As shown in Fig. 17, the presence of TA_6 genotype on both alleles is associated with a higher protein expression while the presence of a TA_7 repeat on only one allele is sufficient to decrease UGT1A1 protein expression. The lowest protein expression level is observed with TA_7 homozygous patients. As shown in Figs. 17b and 17c, the correlative association is also observed between glucuronidaton of the probe substrate estradiol and the number of TA repeats. A similar correlative association is found with SN-38.

As UGT1A1 is considered as a major SN-38 glucuronidation enzyme, we attempted to determine if an association between the expression of this protein and glucuronide formation could exist. As shown in Fig. 18a, there is a positive correlations between glucuronidation of SN-38 and protein level of UGT1A1. To ascertain that the enhancement of glucuronidation observed with this substrate is not attributable to a residual activity of other UGT isoforms, this experiment was reconducted using probe substrates for UGT1A1, namely estradiol,. As seen in fig. 18b a positive correlation exists between UGT1A1 protein level and estradiol-3-G formation. Since estradiol is an endogenously produced

compound and formation of estradiol-3-G is exclusively mediated by UGT1A1, these results demonstrate that a biochemical analysis of serum estradiol-3-G could be properly used to monitor a higher or lower UGT1A1 expression in a patient and therefore, be used as an indicator for determining a predisposition to a physiological reaction to a xenobiotic or an endogenous compound. Finally, Fig. 19 shows the predictive value of the haplotype determination of UGT1A9 and UGT1A1. This haplotype determination includes the genotyping of the *UGT1A9* promoter region and the determination of the number of TA repeats in the TATA box of the *UGT1A1* promoter, which is a more accurate indicator of SN-38 glucuronidation level than the determination of the TA repeats in the TATA box of the *UGT1A1* promoter alone.

EXAMPLE VI

Haplotyping the UGT1A genes

Statistical analysis

Results were expressed as mean ± standard deviation (SD). Differences in kinetic parameters between *UGT* allelic variants were evaluated for statistical significance by paired Student's *t* test. All tests were two-sided. The haplotype frequencies will be estimated using the PHASE 1.0.1 software and Hardy-Weinberg equilibrium and linkage disequilibrium analyses will be performed using ARLEQUIN 2.0™ software.

RESULTS

Analysis of the haplotypic structure of the *UGT1* gene in subjects with UGT1A9*1 or UGT1A9*3 alleles.

Haplotypes of the *UGT1A* gene were analyzed in subjects with the UGT1A9*1/*3 low SN-38 glucuronidation activity genotype.

TABLE 11

UGT1A9 promoter haplotype analysis

Haplotype number	Polymorphisms promoter 1A9 -2208	-2152	-2141	-1887	-1818	-665	-440	-331	-275	-87	1A9 exon 1 codon 33	UGT1A1 TATA box	UGT1A7 codon 208	Haplotype frequency	% of the population
1	C	C	C	T	T	C	T	C	T	G	T			175	31,14
2	C	C	C	T	T	C	T	C	T	A	T			3	0,53
3	C	C	C	T	T	C	C	T	T	G	T			128	22,78
4	C	C	C	T	T	T	C	T	T	G	T			15	2,67
5	C	C	C	T	C	T	C	T	T	G	T			61	10,85
6	C	C	C	G	T	C	C	T	T	G	T			83	14,77
7	C	C	C	G	T	T	C	T	T	G	T			3	0,53
8	C	C	C	T	C	C	C	T	T	G	T			66	11,74
9	T	C	C	T	T	T	C	T	A	G	T			1	0,18
10	C	C	T	T	T	T	C	T	A	G	T			1	0,18
11	C	T	C	T	T	T	C	T	A	G	T			5	0,89
12	C	C	C	T	C	C	C	T	T	G	C			7	1,25
13	C	C	C	G	T	C	T	T	T	G	T			1	0,18
14	C	C	C	T	C	C	T	C	T	G	T			1	0,18
15	C	C	T	G	T	T	C	T	A	G	T			1	0,18
16	C	C	T	G	T	T	C	T	T	G	T			1	0,18
17	C	C	C	T	T	T	C	T	A	G	T			1	0,18
18	C	T	C	T	T	C	C	T	A	G	T			6	1,07
19	C	C	C	T	T	C	C	C	T	G	T			1	0,18
20	C	C	C	T	C	C	C	C	T	G	T			1	0,18
21	C	C	C	T	C	T	C	T	T	G	C			1	0,18

TABLE 12

UGT1A9 and UGT1A1 promoters haplotype analysis

number	-2208	-2152	-2141	-1887	-1818	-665	-440	-331	-275	-87	codon 33	TATA box	codon 208	haplotype frequency	population
22	C	C	C	T	T	T	C	T	T	G	T	6		13	12,04
23	C	C	C	T	T	C	C	T	T	G	T	5		1	0,93
24	C	C	C	T	T	C	T	C	T	G	T	7		23	21,30
25	C	C	C	T	C	T	C	T	T	G	T	6		15	13,89
26	C	C	C	G	T	C	C	T	T	G	T	6		10	9,26
27	C	C	C	T	T	C	T	C	T	G	T	6		3	2,78
28	C	C	C	G	T	T	C	T	T	G	T	7		1	0,93
29	C	C	C	T	T	C	C	T	T	G	T	7		2	1,85
30	C	C	C	T	C	C	C	T	T	G	T	6		13	12,04
31	T	C	C	T	T	T	C	T	A	G	T	6		1	0,93
32	C	C	T	G	T	T	C	T	T	G	T	5		1	0,93
33	C	C	C	T	T	T	C	T	A	G	T	6		1	0,93
34	C	T	C	T	T	T	C	T	A	G	T	7		1	0,93
35	C	T	C	T	T	T	C	T	A	G	T	6		4	3,70
36	C	C	C	T	T	C	C	T	T	G	T	6		2	1,85
37	C	C	C	T	T	C	T	C	T	A	T	7		3	2,78
38	C	C	C	G	T	T	C	T	T	G	T	5		1	0,93
39	C	C	C	T	C	T	C	T	T	G	C	6		1	0,93
40	C	C	C	G	T	T	C	T	T	G	T	6		1	0,93
41	C	C	C	T	C	C	C	T	T	G	C	6		3	2,78
42	C	C	C	T	C	C	C	C	T	G	T	6		1	0,93
43	C	C	C	T	T	C	C	C	T	G	T	7		1	0,93
44	C	T	C	T	T	C	C	T	A	G	T	6		6	5,56

TABLE 13

UGT1A9, UGT1A1 and UGT1A7 haplotype analysis.

Haplotype number	Polymorphisms promoter 1A9 -2208	-2152	-2141	-1887	-1818	-665	-440	-331	-275	-87	1A9 exon 1 codon 33	UGT1A1 TATA box	UGT1A7 codon 208	Haplotype frequency	% of the population
GENBANK [illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
45				T	T	C	T	C			T	7	C	138	30,53
46				T	T	C	C	T			T	6	T	93	20,58
47				G	T	C	C	T			T	6	T	72	15,93
48				T	C	C	C	T			T	6	T	51	11,28
49				T	C	T	C	T			T	6	T	45	9,96
50				T	T	C	C	T			T	6	C	21	4,65
51				T	C	C	C	T			C	6	T	7	1,55
52				T	T	C	T	C			T	6	C	6	1,33
53				T	T	C	C	T			T	7	C	6	1,33
54				T	T	C	C	T			T	7	T	3	0,66
55				T	T	C	T	C			T	7	T	2	0,44
56				T	C	C	C	T			T	7	T	1	0,22
57				G	T	C	T	T			T	6	C	1	0,22
58				T	C	T	C	T			T	6	C	1	0,22
59				G	T	C	C	T			T	6	C	1	0,22
60				T	C	C	T	C			T	6	C	3	0,66
61				T	C	C	C	T			T	6	C	1	0,22
62				T	C	C	C	T			C	7	C	1	0,22

TABLE 14

Allele Frequencies

	-2208 Allele or genotype	-2208 Frequency Am	-2152 Allele or genotype	-2152 Frequency Am	-2141 Allele or genotype	-2141 Frequency Am	-1887 Allele or genotype	-1887 Frequency Am
	C	0,99	C	0,95	C	0,99	T	0,85
	T	0,01	T	0,05	T	0,01	G	0,15
Homo. WT	CC	0,98	CC	0,90	CC	0,98	TT	0,75
Heterozygous	CT	0,02	CT	0,10	CT	0,02	TG	0,21
Homo. var.	TT	0,00	TT	0,00	TT	0,00	GG	0,04
		N=48		N=48		N=48		N=48

-1818 Allele or genotype	-1818 Frequency Am	-665 Allele or genotype	-665 Frequency Am	-440 Allele or genotype	-440 Frequency Am	-331 Allele or genotype	-331 Frequency Am	-275 Allele or genotype	-275 Frequency Am
T	0,71	C	0,58	T	0,30	C	0,30	T	0,92
C	0,29	T	0,42	C	0,70	T	0,70	A	0,08
TT	0,50	CC	0,23	TT	0,15	CC	0,15	TT	0,85
TC	0,42	CT	0,71	TC	0,31	CT	0,31	TA	0,15
CC	0,08	TT	0,06	CC	0,54	TT	0,54	AA	0,00
	N=48		N=48		N=48		N=48		N=46

TABLE 15

Functional UGT1A1, UGT1A7 and UGT1A9 SNPs frenquence in the French-Canadian population.

	UGT1A9 Codon 33		UGT1A7 Codon 208		UGT1A1 TATA box	
Wild-type allele	T	0,98	T	0,62	6	0,67
Mutant allele	C	0,02	C	0,38	7	0,33

EXAMPLE VII

Multiple protein sequence alignment of UGT1A proteins at selected positions

UGT1A7*1, UGT1A9*1 and their genetic variant proteins UGT1A7 (a) and UGT1A9 (b) are aligned with close members of the UGT1A subfamily and the rat UGT1A7 isoenzyme. The varying amino acid positions are indicated with bold characters.

DISCUSSION

After resequencing the first exons of *UGT1A7* and *UGT1A9* genes, 4 polymorphic sites in the targeted regions were identified. Two polymorphic UGT1A9 variants were discovered, UGT1A9*2 $C^{3}Y$ and UGT1A9*3 $M^{33}T$. In addition, the presence of two novel nonsynonymous UGT1A7 SNPs, $G^{115}S$ and $E^{139}D$, combined with previously described missense polymorphisms at codons 129/131 and 208, generated five additional UGT1A7 alleles (*5 through *9). Based on the *in vitro* functional genomic assays, the UGT1A7*3, *4, *5, *8 and *9 alleles and the UGT1A9*3 allele were all identified as low SN-38 glucuronidating alleles. Results demonstrate that the coinheritance of UGT1A1, UGT1A7 variants and especially the loss of function UGT1A9 polymorphism determine individual's susceptibility to irinotecan-induced toxicity. Thus, findings lay emphasis on the necessity to analyze combination of UGT1A1, UGT1A7 and UGT1A9 polymorphisms (haplotypes) rather than looking for a single

polymorphism present in the *UGT1A1* gene to predict patients at higher risk of developing irinotecan-induced toxicity in a clinical setting.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth, and as follows in the scope of the appended claims.

CLAIMS:

1. A method for determining predisposition to a physiological reaction of an individual to a biologically active compound comprising characterizing nucleotide sequence of at least one of the *UGT1A1*, *UGT1A7* or *UGT1A9* gene or a part thereof of said individual, wherein the presence of at least one polymorphic or haplotypic variation in said nucleotide sequence is indicative of said predisposition to a physiological reaction.

2. The method of claim 1, wherein said predisposition is a hereditary predisposition.

3. The method of claim 1, wherein said predisposition is a higher or lower susceptibility, sensibility, diathesis, proneness, proclivity, tendency, sensitivity, responsiveness, resistance or constitutional sickness to said physiological reaction.

4. The method of claim 1, wherein said physiological reaction is a beneficial reaction.

5. The method of claim 1, wherein said physiological reaction is an adverse reaction or a side effect.

6. The method of claim 1, wherein said biologically active compound is a xenobiotic.

7. The method of claim 6, wherein said xenobiotic is a drug, a carcinogen or a pre-carcinogen.

8. The method of claim 7, wherein said drug is an anti-cancer agent or an immunosuppressive agent.

9. The method of claim 8, wherein said anti-cancer agent is a camptothecin or an analog thereof.

10. The method of claim 9, wherein said camptothecin analog is 7-ethyl-10-[4-(1-piperidino)-1-piperidino] carbonyloxy camptothecin (irinotecan, CPT-11), 7-ethyl-10-hydroxycamptothecin (SN-38).

11. The method of claim 8, wherein said immunosuppressive agent is mycophenolic acid (MPA).

12. The method of claim 1, wherein said individual is a human or an animal.

13. The method of claim 1, wherein said individual is a patient with cancer.

14. The method of claim 13, wherein said patient has a colorectal cancer or a solid tumor.

15. The method of claim 1, wherein determining genetic sequence is performed on a DNA or a RNA sample.

16. The method of claim 1, wherein said polymorphic or haplotypic variation is a UGT1A9 variation.

17. The method of claim 16, wherein said UGT1A9 variation is at least one of a $C^{-2208}T$ substitution, a $C^{-2152}T$ substitution, a $C^{-2141}T$ substitution, a $T^{1887}G$ substitution, a $T^{-1818}C$ substitution, a $C^{-665}T$ substitution, a $T^{-440}C$ substitution, a $C^{-331}T$ substitution, a $T^{-275}A$ substitution, a $G^{-87}A$ substitution, a $G^{8}A$ missence mutation ($C^{3}Y$), a $T^{98}C$ missence mutation ($M^{33}T$) or combination thereof.

18. The method of claim 17, wherein said $G^{8}A$ missence mutation is associated with a decreased predisposition or susceptibility to an anti-cancer agent.

19. The method of claim 17, wherein said $G^{8}A$ missence mutation is associated with a decreased responsiveness to an immunosuppressive agent.

20. The method of claim 17, wherein said $T^{98}C$ missence mutation is associated with an increased adverse reaction to an anti-cancer agent.

21. The method of claim 1, wherein said polymorphic or haplotypic variation is a UGT1A7 variation.

22. The method of claim 21, wherein said UGT1A7 variation is a $G^{353}T$ missense mutation, a $T^{397}G$ missense mutation, a $C^{401}A$ missense mutation, a $G^{402}A$ missense mutations, a $G^{427}C$ missense mutation, a $T^{632}C$ missense mutation or combination thereof.

23. The method of claim 1, wherein said polymorphic or haplotypic variation is a UGT1A1 variation.

24. The method of claim 23, wherein said UGT1A1 variation is a TA_7 mutation in the TATA box.

25. An isolated nucleotide sequence comprising at least one nucleotide sequence selected from the group consisting of SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45 SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, SEQ ID NO: 60, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67, SEQ ID NO: 68, a fragment or the complementary sequences thereof, for determining predisposition to a physiological reaction.

26. The nucleotide sequence of claim 25, wherein said sequence is an allelic variant of UGT1A1, UGT1A7 or UGT1A9.

27. An isolated amino acid sequence comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO: 69, SEQ ID NO: 70, SEQ ID NO: 71 or a fragment thereof.

28. The amino acid sequence of claim 27, wherein said sequence is encoded by a nucleotide sequence comprising at least one sequence selected from the group consisting of SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, a fragment or the complementary sequences thereof.

29. The amino acid sequence of claim 27, wherein the expression of said sequence is regulated by a nucleotide sequence comprising at least one sequence selected from the group consisting of SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 42, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45 SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 48, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, a fragment or the complementary sequences thereof.

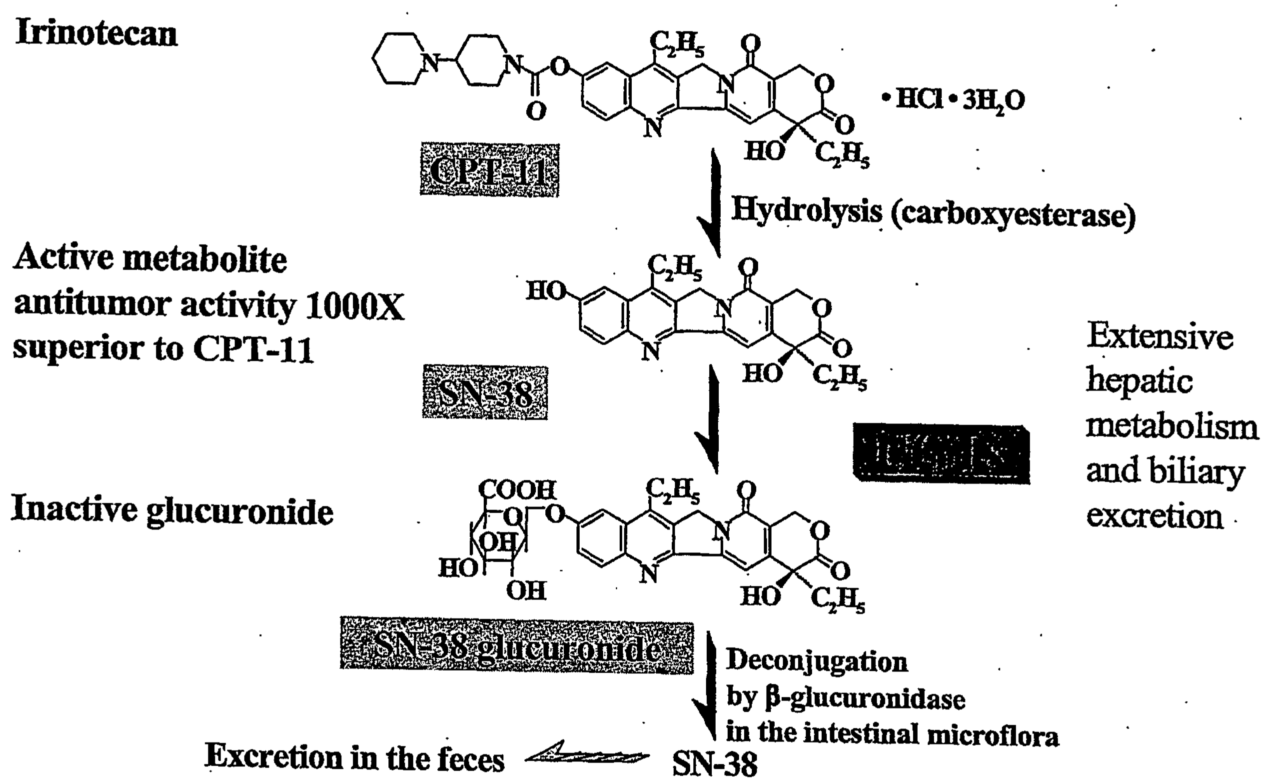

Irinotecan
•HCl•3H_2O
CPT-11
Hydrolysis (carboxyesterase)
Active metabolite
antitumor activity 1000X
superior to CPT-11
SN-38
Extensive
hepatic
metabolism
and biliary
excretion
Inactive glucuronide
SN-38 glucuronide
Deconjugation
by β-glucuronidase
in the intestinal microflora
Excretion in the feces
SN-38

Fig. 1

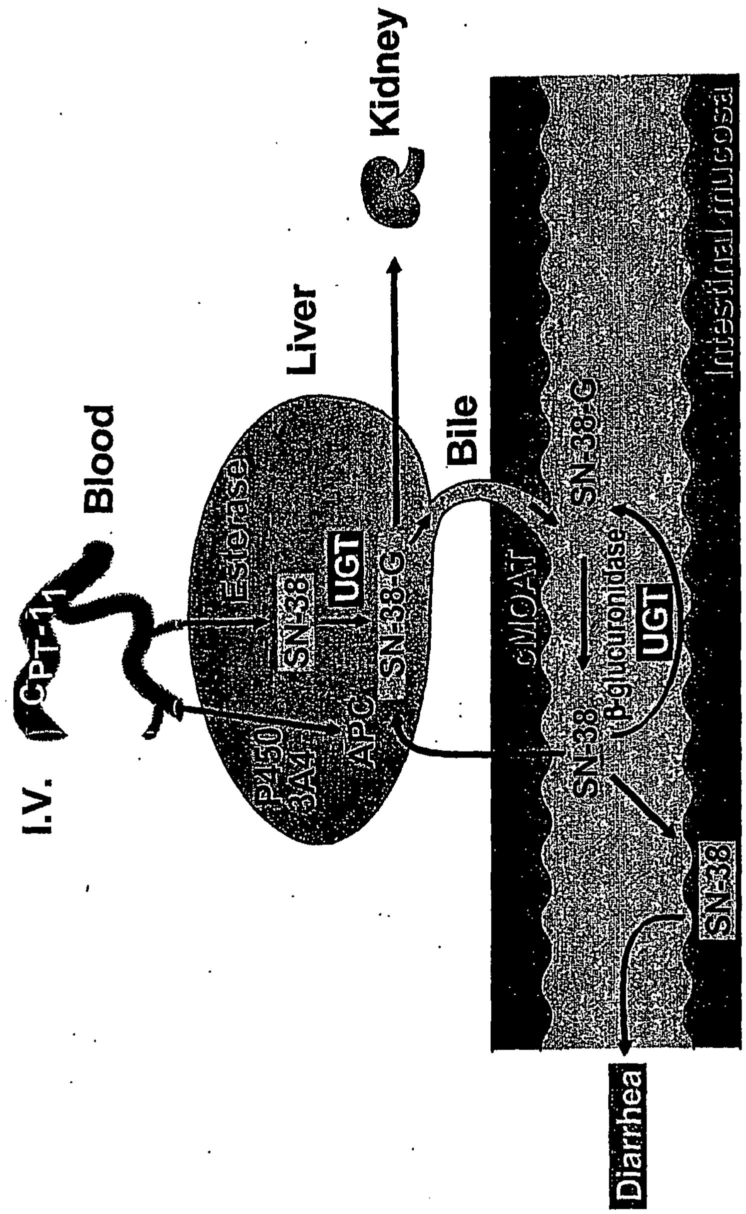

I.V.
CPT-11
Blood
P450
3A4
APC
Esterase
SN-38
UGT
SN-38-G
Liver
Kidney
Bile
cMOAT
SN-38
β-glucuronidase
SN-38-G
UGT
SN-38
Intestinal mucosa
Diarrhea

Fig. 2

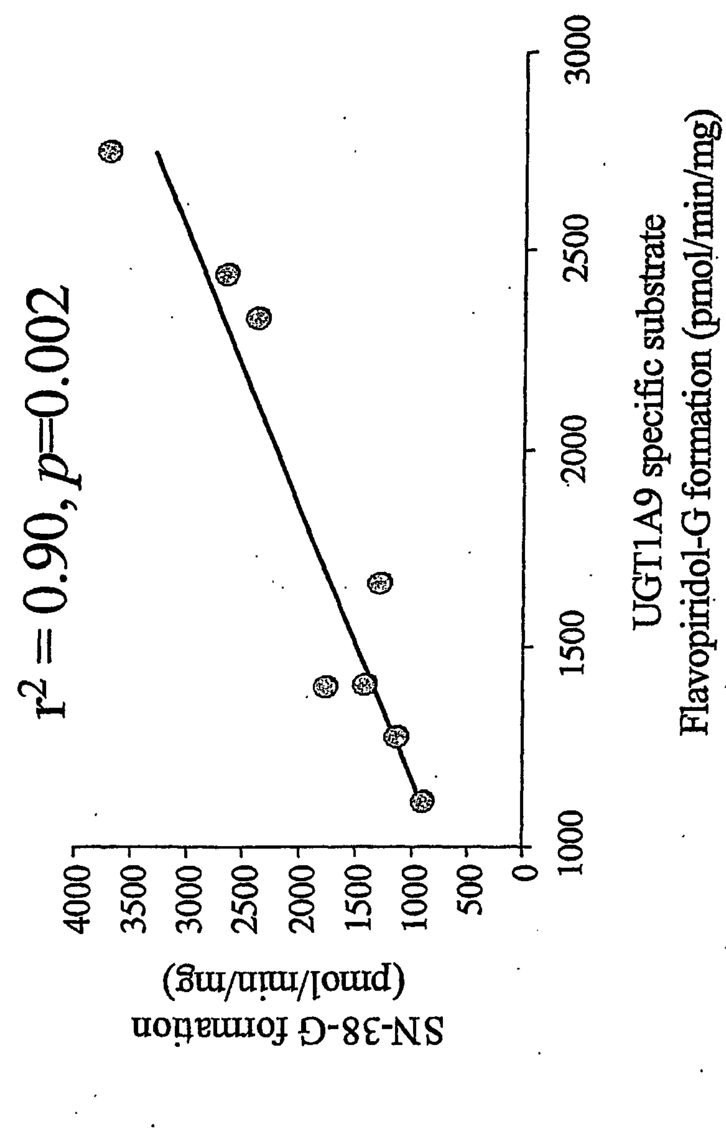

r2 = 0.90, p=0.002
SN-38-G formation
(pmol/min/mg)
4000
3500
3000
2500
2000
1500
1000
500
0
1000
1500
2000
2500
3000
UGT1A9 specific substrate
Flavopiridol-G formation (pmol/min/mg)

Fig. 3

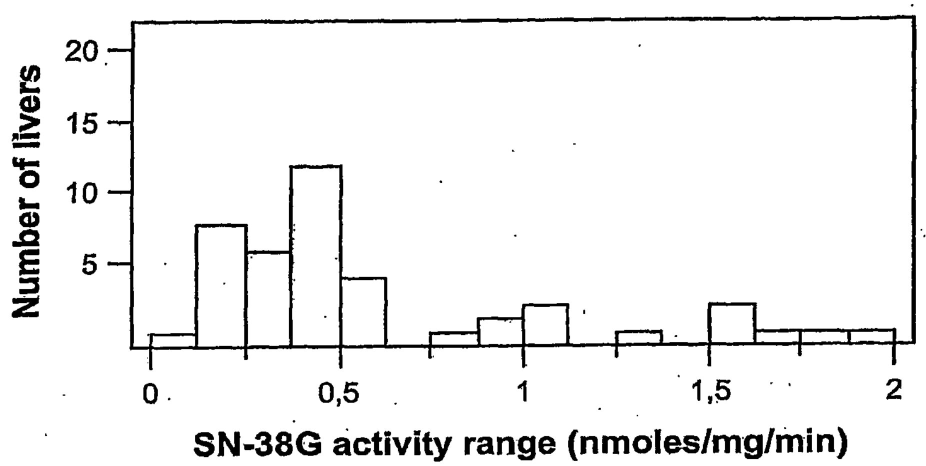

Number of livers
20
15
10
5
0
0,5
1
1,5
2
SN-38G activity range (nmoles/mg/min)

Fig. 4

5/28

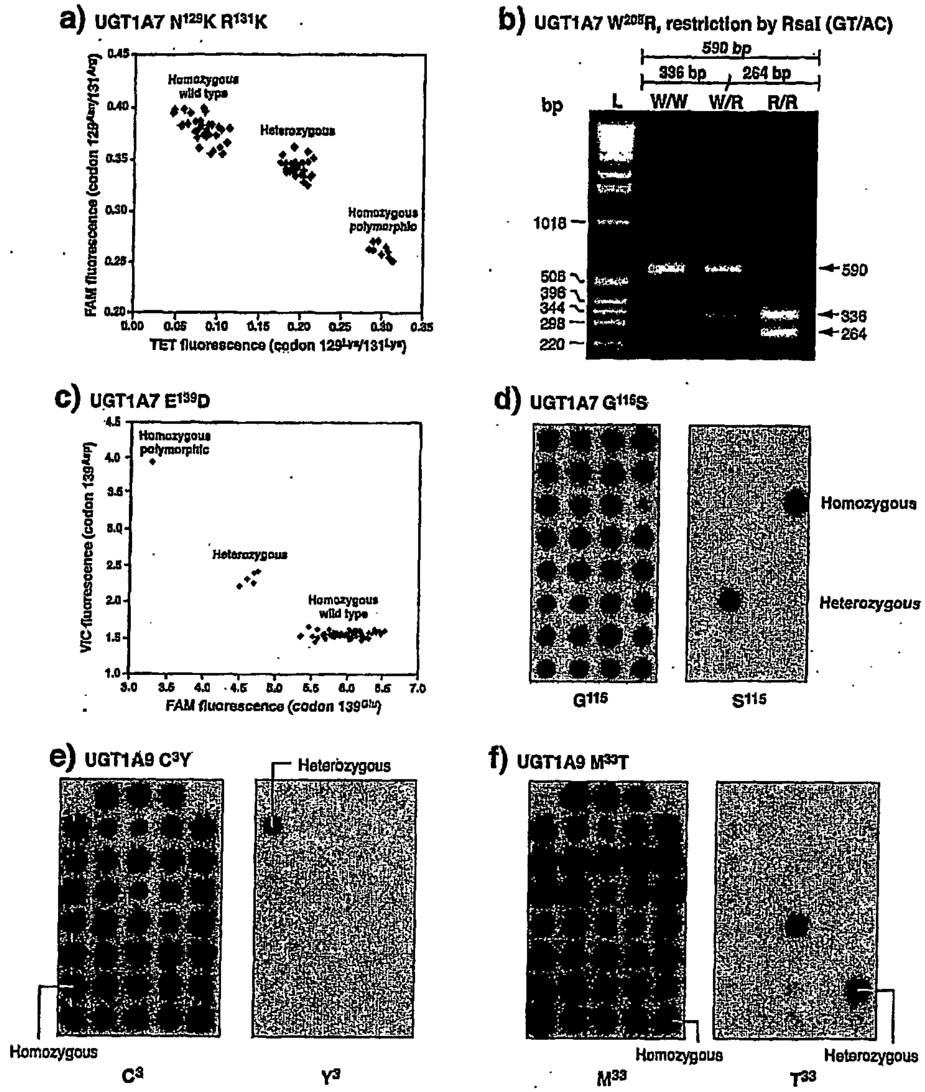

a) UGT1A7 N129K R131K
Homozygous wild type
Heterozygous
Homozygous polymorphic
FAM fluorescence (codon 129Asn/131Arg)
TET fluorescence (codon 129Lys/131Lys)
b) UGT1A7 W208R, restriction by RsaI (GT/AC)
590 bp
336 bp
264 bp
bp
L
W/W
W/R
R/R
1018
506
396
344
298
220
590
336
264
c) UGT1A7 E139D
Homozygous polymorphic
Heterozygous
Homozygous wild type
VIC fluorescence (codon 139Asp)
FAM fluorescence (codon 139Glu)
d) UGT1A7 G115S
Homozygous
Heterozygous
G115
S115
e) UGT1A9 C3Y
Heterozygous
Homozygous
C3
Y3
f) UGT1A9 M33T
Homozygous
Heterozygous
M33
T33

Fig. 5

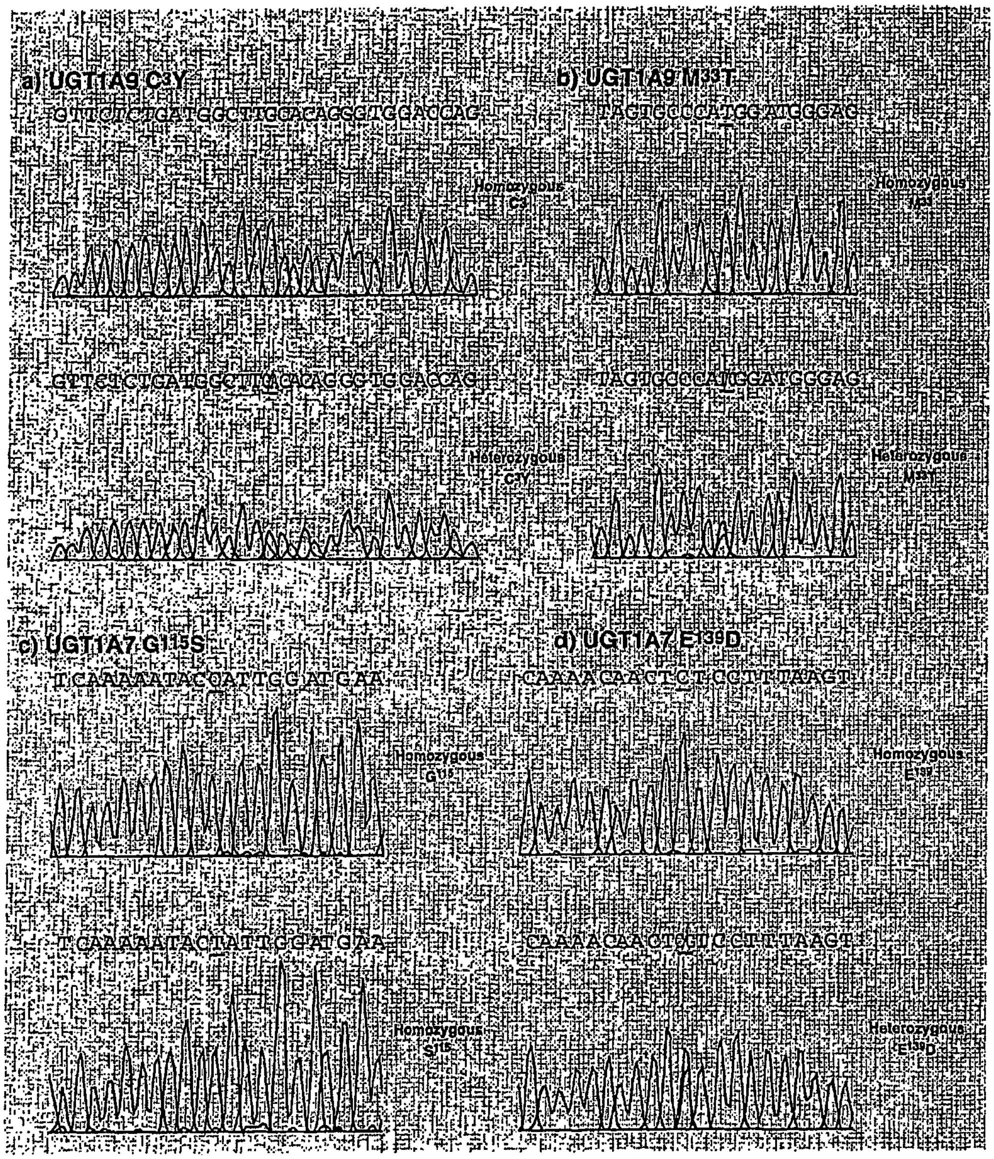

a) UGT1A9 C3Y
b) UGT1A9 M33T
Homozygous
Heterozygous
c) UGT1A7 G115S
d) UGT1A7 E139D
Homozygous
Heterozygous

Fig. 6

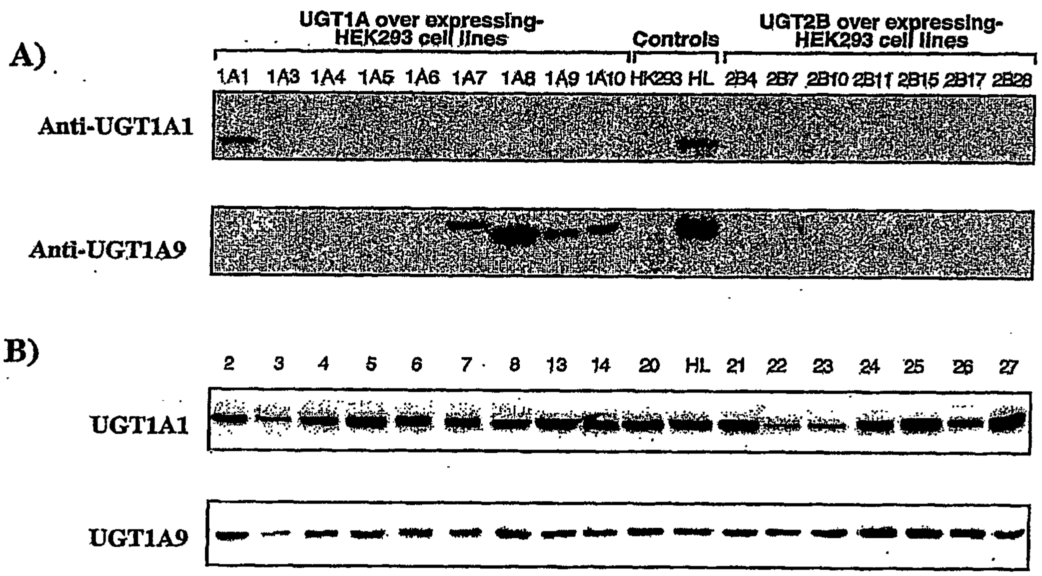
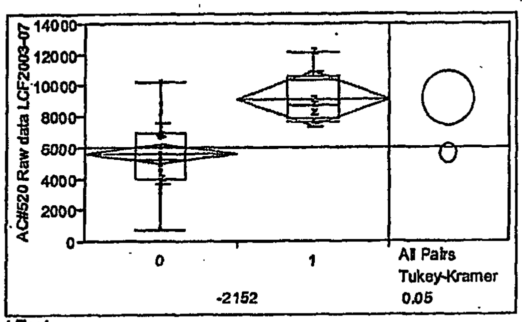

A)
UGT1A over expressing-HEK293 cell lines
Controls
UGT2B over expressing-HEK293 cell lines
1A1 1A3 1A4 1A5 1A6 1A7 1A8 1A9 1A10 HK293 HL 2B4 2B7 2B10 2B11 2B15 2B17 2B28
Anti-UGT1A1
Anti-UGT1A9
B)
2 3 4 5 6 7 8 13 14 20 HL 21 22 23 24 25 26 27
UGT1A1
UGT1A9

Fig. 7

Oneway Analysis of UGT1A9 protein expression By promoter variant

-2152 non-carrier: (0)
-2152 carrier: (1)



t-Test

	Difference	t-Test	DF	Prob > |t|
Estimate	-3511.44	-3.805	46	0.0004

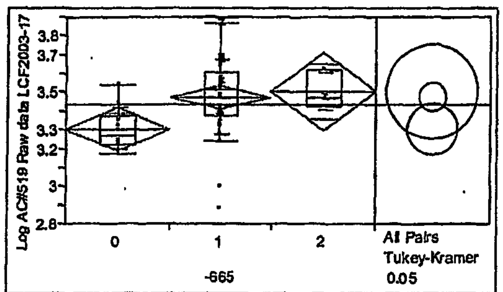
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
-2152	1	55229128	55229128	14.4795	0.0004

Fig. 8a

Oneway Analysis of UGT1A9 protein expression By –665 promoter variant

-665 non carrier: (0)
-665 heterozygous: (1)
-665 homozygous: (2)



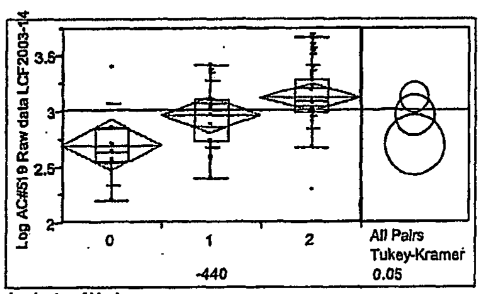
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
-665	2	0.2354303	0.117715	3.6937	0.0331

Fig. 8b

Oneway Analysis of UGT1A9 protein expression By -440 promoter variant

-440 non carrier: (0)
-440 heterozygous: (1)
-440 homozygous: (2)

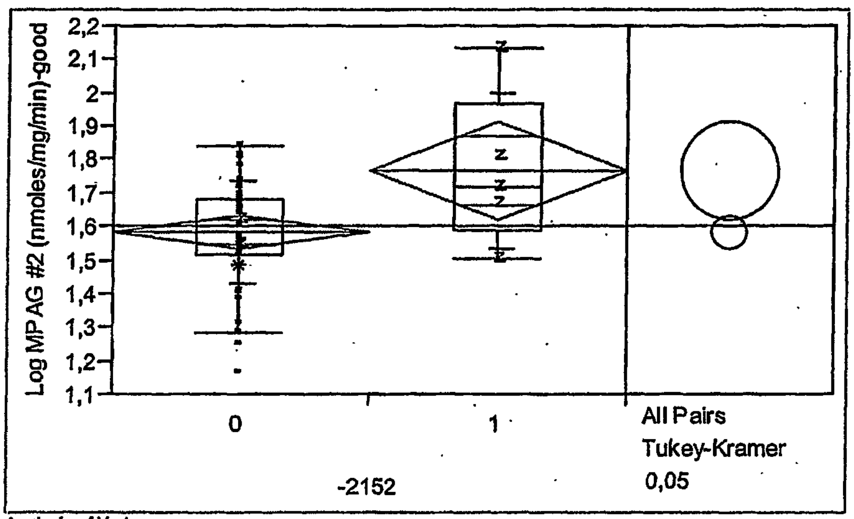


Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
-440	2	1.0529458	0.526473	5.8967	0.0053

Fig. 8c

-2152 carriers: (1)
-2152 non carriers: (O)



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
-2152	1	0,1486778	0,148678	5,6323	0,0219
Error	46	1,2142877	0,026398		

Fig. 9

–1818 non carriers: (0)
–1818 heterozygous: (1)
–1818 homozygous: (2)

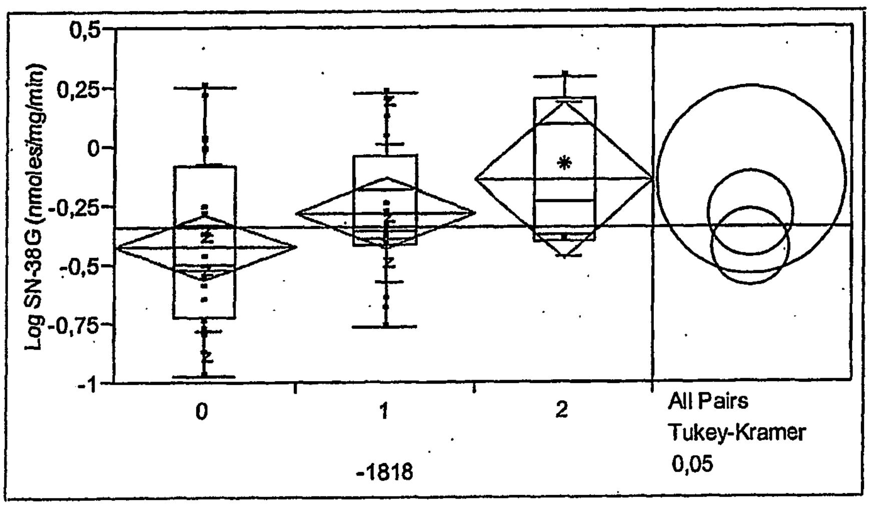


Fig. 10

SN-38-Glucuronide formation

-665 carriers: (1)
-665 non carriers: (0)

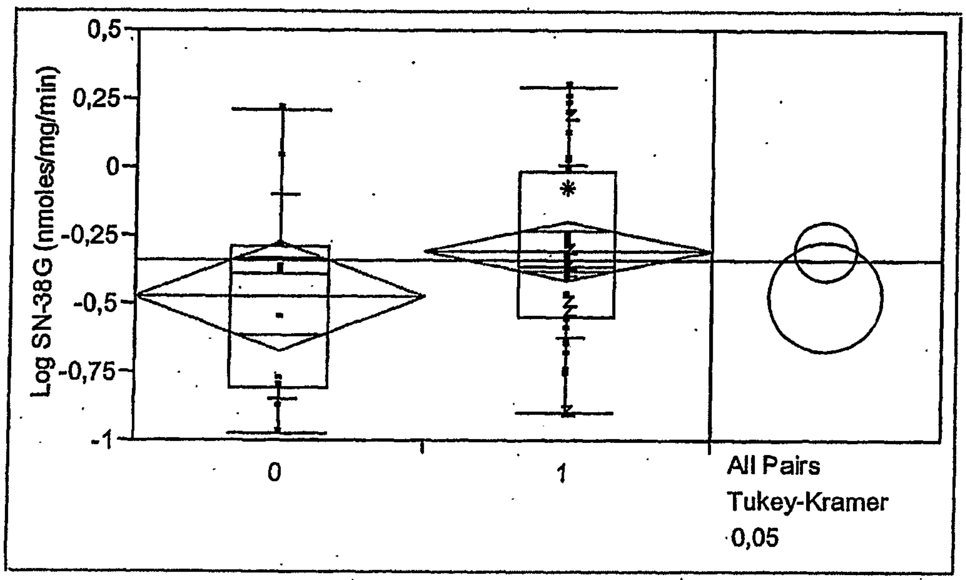
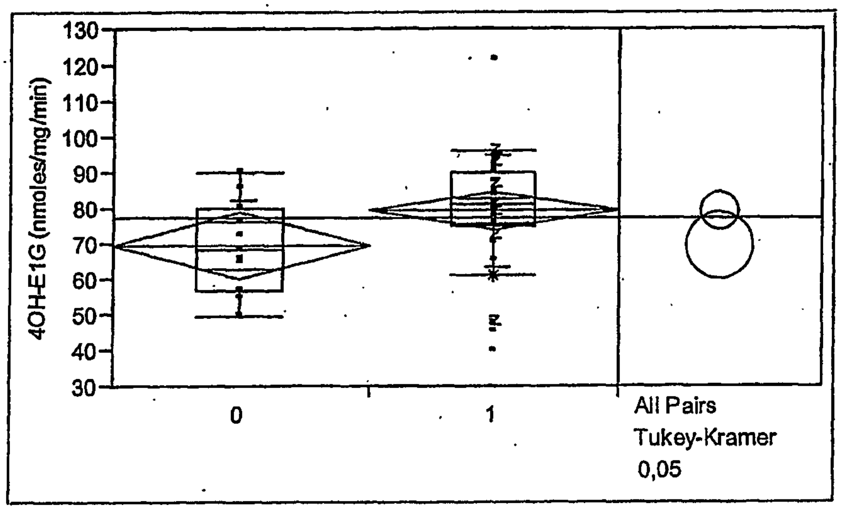


Fig. 11a

4OHEstrone-Glucuronide formation

-665 carriers: (1)
-665 non carriers: (0)



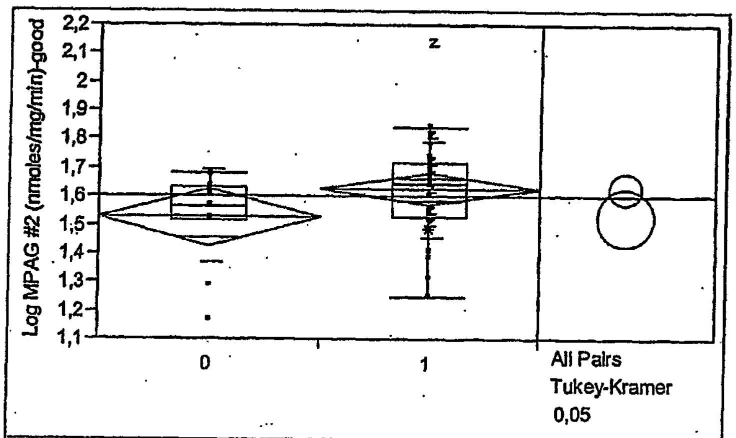
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Porteur vs non porteur de mut -665	1	822,270	822,270	3,4737	0,0687
Error	46	10888,756	236,712		
C. Total	47	11711,026			

Fig. 11b

MPA-glucuronide formation

-665 carriers: (1)
-665 non carriers: (0)

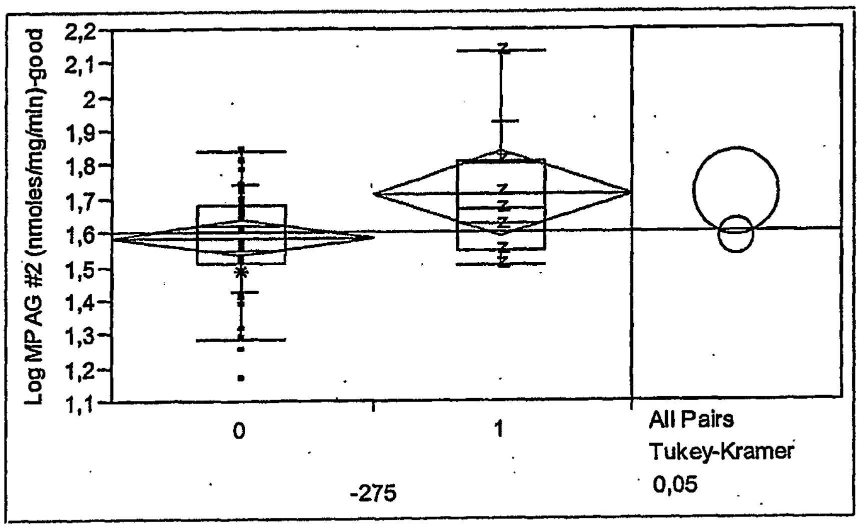


Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Porteur vs non porteur de mut -665	1	0,0800407	0,080041	2,8698	0,0970
Error	46	1,2829249	0,027890		
C. Total	47	1,3629656			

Fig. 11c

-275 carriers: (1)
-275 non carriers: (0)

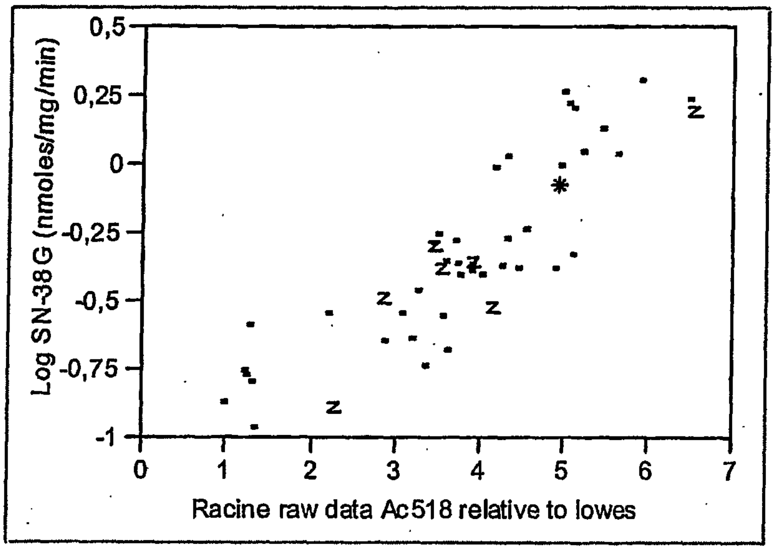


Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
-275	1	0,1011341	0,101134	3,6868	0,0611
Error	46	1,2618314	0,027431		
C. Total	47	1,3629656			

Fig. 12

SN-38 G formation (nmoles/mg/min) By UGT1A1 protein level

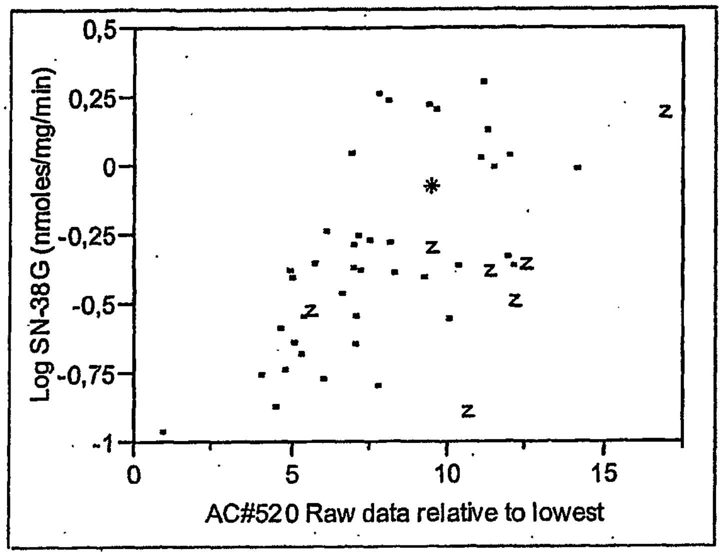


Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-1,153565	0,071888	-16,05	<.0001
Racine raw data Ac518 relative to lowes	0,2128557	0,017751	11,99	<.0001

Fig. 13a

SN-38 G formation (nmoles/mg/min) By UGT1A9 protein level (relative to the lowest)



Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0,853447	0,117342	-7,27	<.0001
AC#520 Raw data relative to lowest	0,0614746	0,013224	4,65	<.0001

Fig. 13b

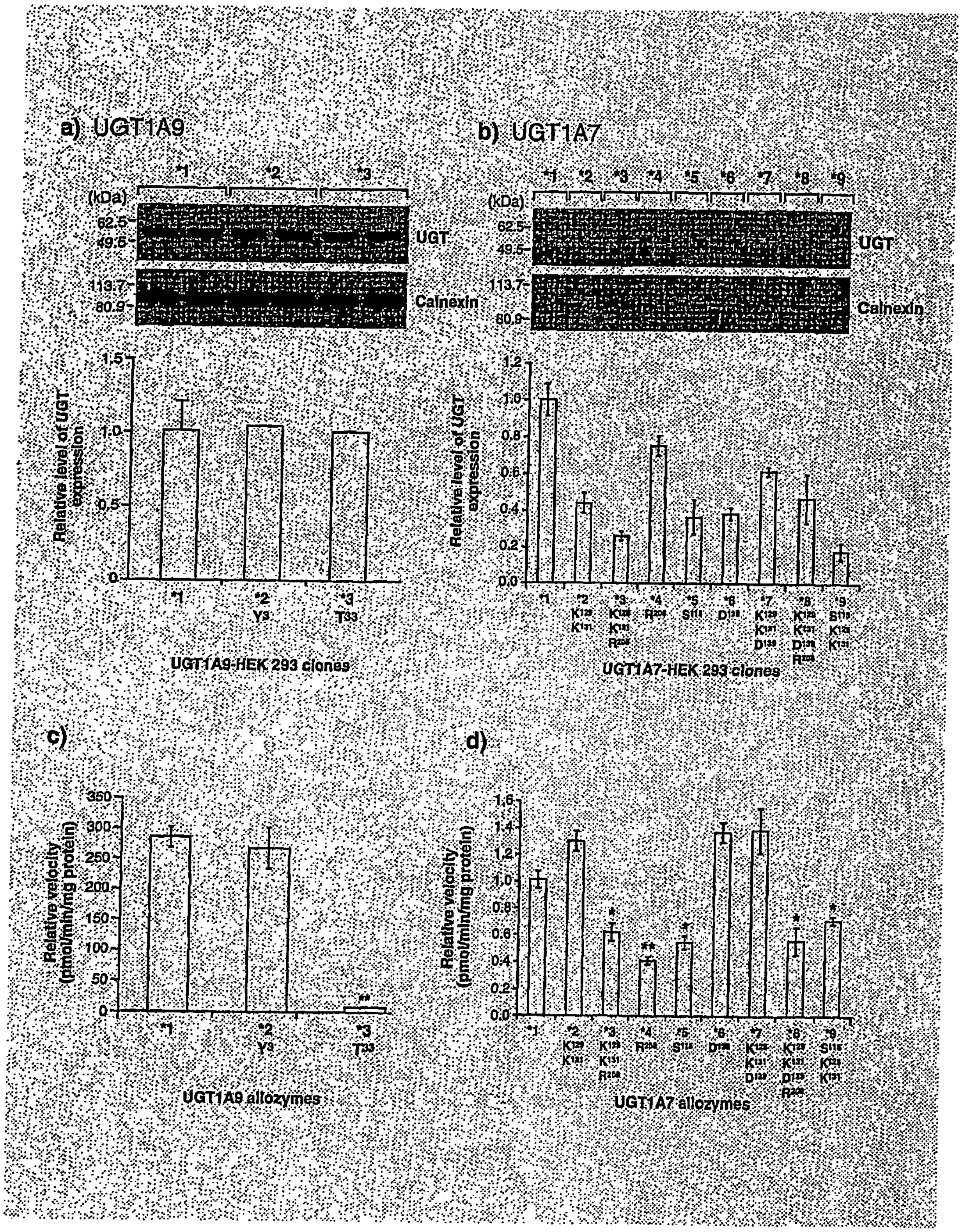

a) UGT1A9
b) UGT1A7
(kDa)
UGT
Calnexin
Relative level of UGT expression
UGT1A9-HEK 293 clones
UGT1A7-HEK 293 clones
c)
d)
Relative velocity (pmol/min/mg protein)
UGT1A9 allozymes
UGT1A7 allozymes

Fig. 14

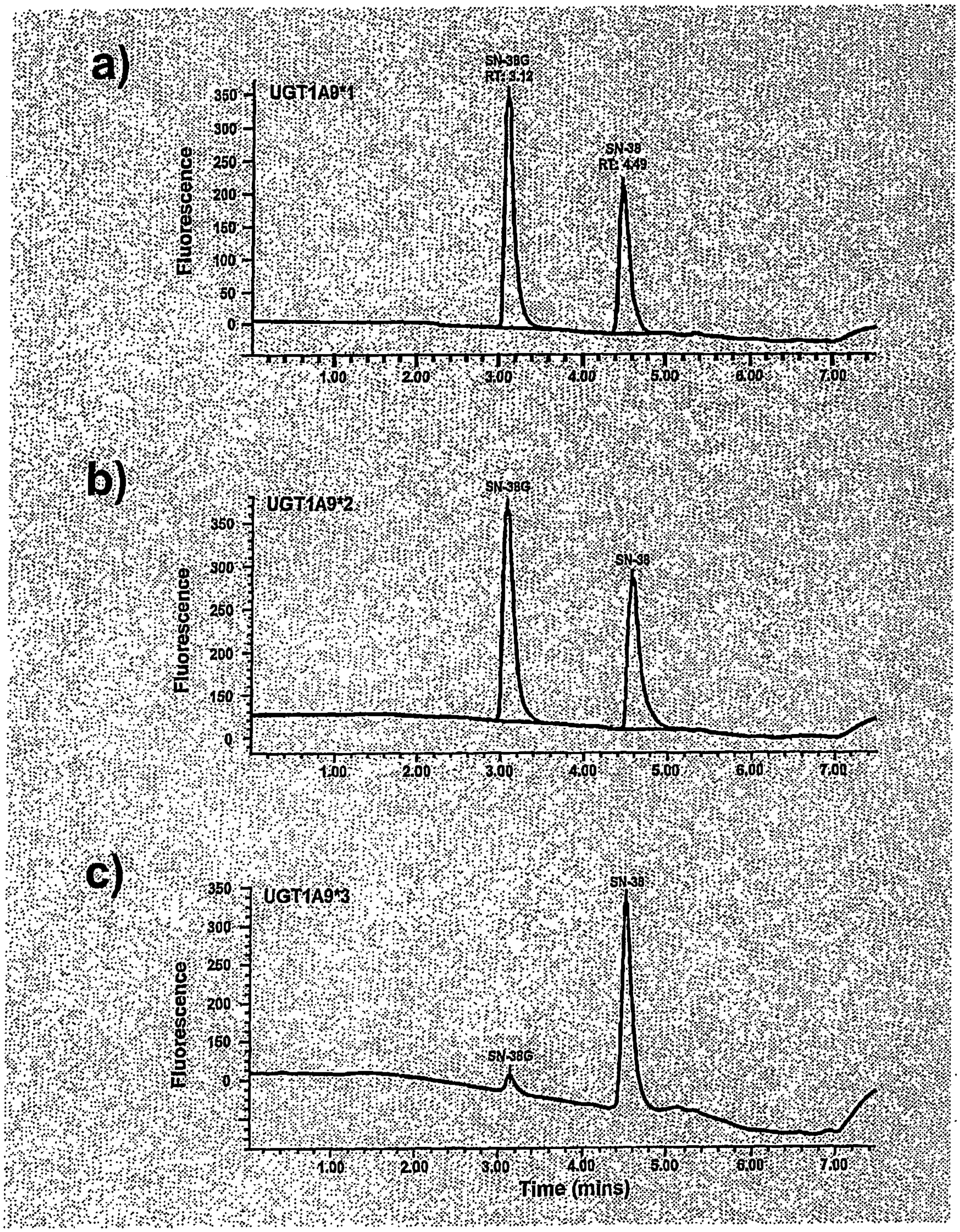

a)
UGT1A9*1
SN-38G
RT: 3.12
SN-38
RT: 4.49
Fluorescence
b)
UGT1A9*2
SN-38G
SN-38
Fluorescence
c)
UGT1A9*3
SN-38
SN-38G
Fluorescence
Time (mins)

Fig. 15

Fig. 16

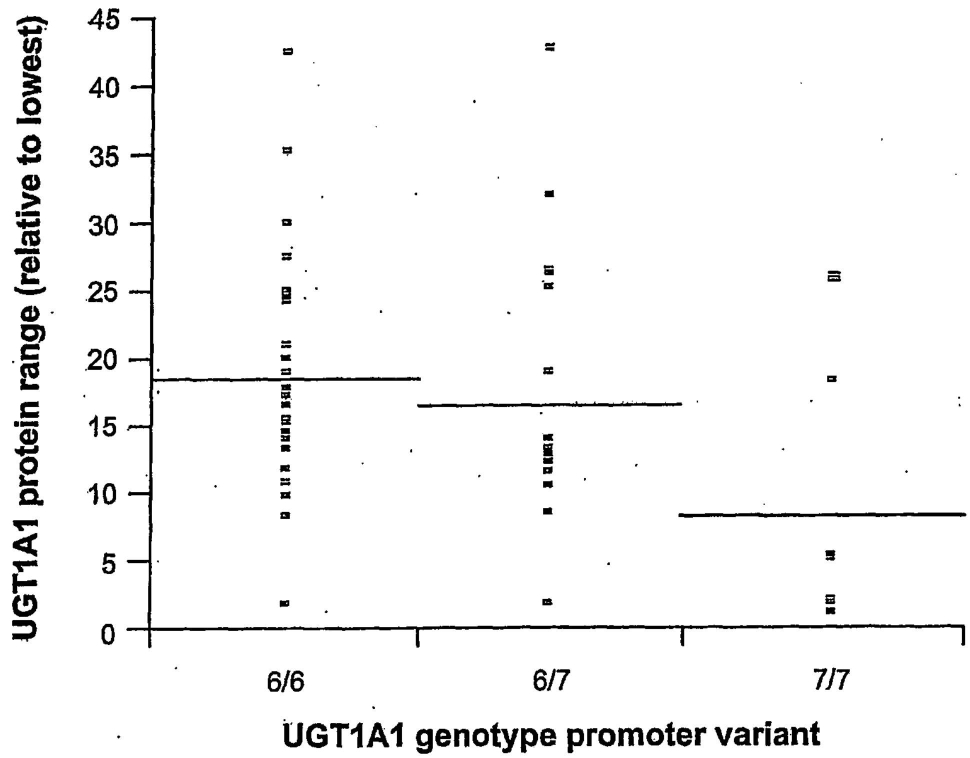

UGT1A1 protein range (relative to lowest)
45
40
35
30
25
20
15
10
5
0
6/6
6/7
7/7
UGT1A1 genotype promoter variant

Fig. 17a

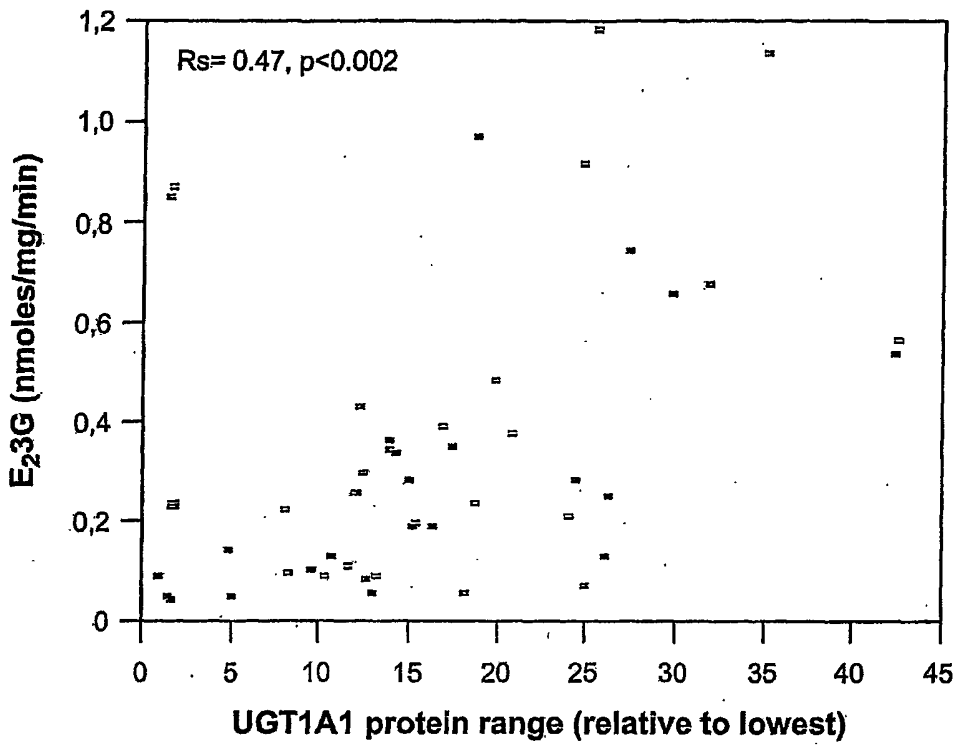
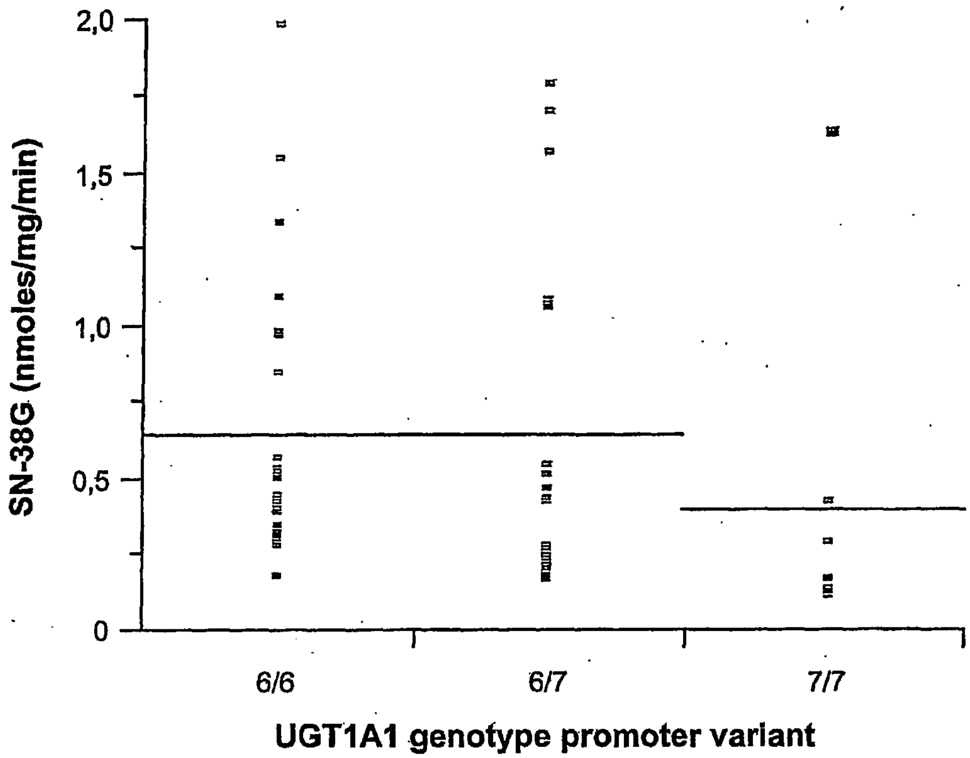
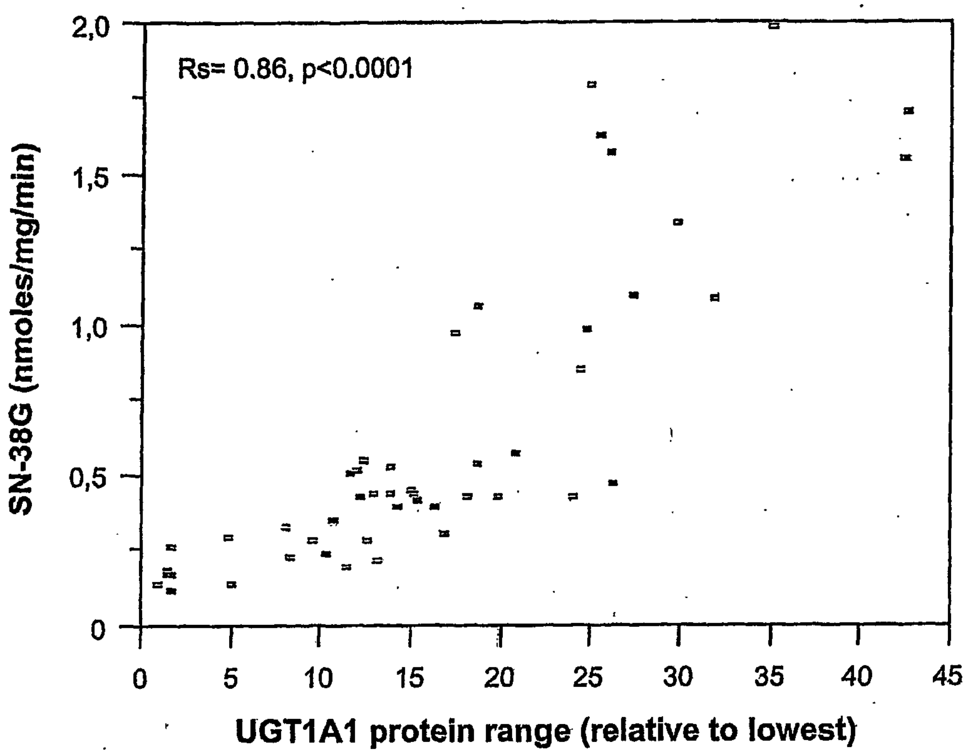

Rs= 0.47, p<0.002
1,2
1,0
0,8
0,6
0,4
0,2
0
E23G (nmoles/mg/min)
0
5
10
15
20
25
30
35
40
45
UGT1A1 protein range (relative to lowest)

Fig. 17b



SN-38G (nmoles/mg/min)
2,0
1,5
1,0
0,5
0
6/6
6/7
7/7
UGT1A1 genotype promoter variant

Fig. 17c



Rs= 0.86, p<0.0001
SN-38G (nmoles/mg/min)
2,0
1,5
1,0
0,5
0
0
5
10
15
20
25
30
35
40
45
UGT1A1 protein range (relative to lowest)

Fig. 18a

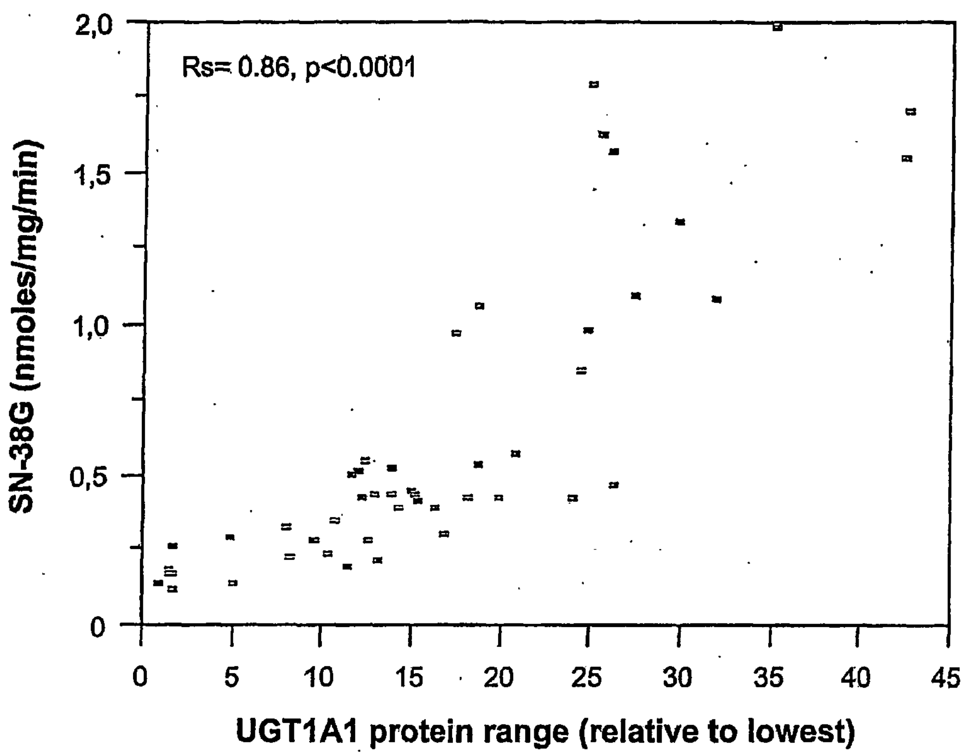

Rs= 0.86, p<0.0001
SN-38G (nmoles/mg/min)
2,0
1,5
1,0
0,5
0
0
5
10
15
20
25
30
35
40
45
UGT1A1 protein range (relative to lowest)

Fig. 18b

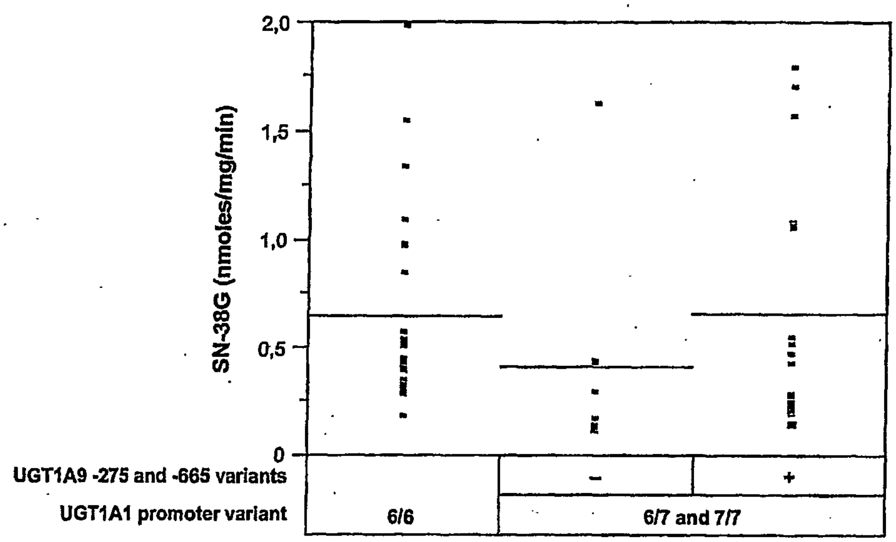

2,0
1,5
1,0
0,5
0
SN-38G (nmoles/mg/min)
UGT1A9 -275 and -665 variants
UGT1A1 promoter variant
–
+
6/6
6/7 and 7/7

Fig. 19

a)

	115 129 131 139	208
UGT1A7	LLTSSSNGIFDLFFSNCRSLFNDRKLVEYLKESCFDAVF	MTFKERVWNHIMHLE
UGT1A7V	LLTSSSNSIFDLFFSNCRSLFKDKKLVEYLKDSCFDAVF	MTFKERVRNHIMHLE
UGT1A8	LFLSSSNGFFNLFFSHCRSLFNDRKLVEYLKESSFDAVF	MTFKERVRNHIMHLE
UGT1A9	LLMGSYNDIFDLFFSNCRSLFKDKKLVEYLKESSFDAVF	MTFKERVRNHIMHLE
UGTA10	LLMSSSSGFLDLFFSHCRSLFNDRKLVEYLKESSFDAVF	MTFKERVWNHIVHLE
UGT1A7Rat	LLTSPAQGFFELLFSHCRSLFKDKKLVEYLKQSSFDAVF	MTFKERVWNLLSYMG

b)

	3	33
UGT1A9	MACTGW	VVPMDGS
UGT1A9V	MAYTGW	VVPTDGS
UGT1A1	MAVESQ	LIPVDGS
UGT1A3	MATGLQ	LVPIDGS
UGT1A4	MARGLQ	LVPTDGS
UGT1A5	MATGLQ	LVPTDGS
UGT1A6	MACLLR	LVPQDGS
UGT1A7	MARAGW	LVPMDGS
UGT1A8	MARTGW	LVPMDGS
UGT1A10	MARAGW	LVPMDGS

Fig. 20

SEQUENCE LISTING

<110> UNIVERSITE LAVAL
Guillemette, Chantal

<120> Method for determining predisposition to
a physiological reaction in a patient

<130> 6013-118PCT

<150> 60/412,002
<151> 2002-09-20

<160> 71

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(17)
<223> UGT1A9 #37 (Forward)

<400> 1
gtgctggtat ttctccc 17

<210> 2
<211> 24
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(24)
<223> UGT1A9 #38 (Reverse)

<400> 2
gtcaaaaatg tcattgtatg aacc 24

<210> 3
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(20)
<223> UGT1A9 #39 (Forward)

<400> 3
gatctggacc gggagttcaa 20

<210> 4
<211> 22
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(22)
<223> UGT1A9 #40 (Reverse)

<400> 4
gtgtggctgt agagatcata ct 22

<210> 5
<211> 25
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(25)
<223> UGT1A9 #41 (Forward)

<400> 5
catgcacttg gaggaacatt tatta 25

<210> 6
<211> 18
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(18)
<223> UGT1A9 #42 (Reverse)

<400> 6
gagtacacgc attggcac 18

<210> 7
<211> 18
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(18)
<223> UGT1A7 #18 (Forward)

<400> 7
cgctggacgg caccattg 18

<210> 8
<211> 22
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(22)
<223> UGT1A7 #17 (Reverse)

<400> 8
gctaaagggg agataactta cc 22

<210> 9
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(17)
<223> UGT1A7 #122 (Forward)

<400> 9
gctggacggc accattg 17

<210> 10
<211> 19
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(19)
<223> UGT1A7 #123 (Reverse)

<400> 10
ccctaagaga agtctgggg 19

<210> 11
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(17)
<223> UGT1A9 #7 (Forward)

<400> 11
ctcccaccta ctgtatc 17

<210> 12
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(17)
<223> UGT1A9 #8 (Forward)

<400> 12
gttcaaggct tttgccc 17

<210> 13
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(17)
<223> UGT1A9 #9 (Forward)

<400> 13
catttattat gccaccg 17

<210> 14
<211> 16
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(16)
<223> ASO UGT1A9 C3 (Forward)

<400> 14
atggcttgca cagggt 16

<210> 15
<211> 16
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(16)
<223> ASO UGT1A9 Y3 (Forward)

<400> 15
atggcttaca cagggt 16

<210> 16
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(17)
<223> ASO UGT1A9 M33 (Forward)

<400> 16
agtgcccatg gatggga 17

<210> 17

<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(19)
<223> ASO UGT1A9 T33 (Forward)

<400> 17
agtgcccacg gatggga 17

<210> 18
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(17)
<223> ASO UGT1A7 G115 (Forward)

<400> 18
catccaatgg tattttt 17

<210> 19
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(17)
<223> ASO UGT1A7 S115 (Forward)

<400> 19
catccaatag tattttt 17

<210> 20
<211> 19
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(19)
<223> Taqman UGT1A7 codon 139/131 #387 (Forward)

<400> 20
gcaccattgc gaagtgcat 19

<210> 21
<211> 22
<212> DNA
<213> Homo sapiens

<220>

<221> primer_bind
<222> (1)...(22)
<223> Taqman UGT1A7 codon 139/131 #388 (Reverse)

<400> 21
ggatcgagaa acactgcatc aa 22

<210> 22
<211> 16
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(16)
<223> Taqman UGT1A7 codon 139/131 K129/K131-FAM
(Forward)

<400> 22
ttaatgaccg aaaatt 16

<210> 23
<211> 17
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(17)
<223> Taqman UGT1A7 codon 139/131 K129/K131-TET
(Forward)

<400> 23
tttaaggaca aaaaatt 17

<210> 24
<211> 26
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(26)
<223> Taqman UGT1A7 codon 139 #546 (Forward)

<400> 24
gcgaagtgca ttttctctat taacaa 26

<210> 25
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(20)
<223> Taqman UGT1A7 codon 139 #544 (Reverse)

<400> 25
aagccacagc gatcaaaagg 20

<210> 26
<211> 21
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(21)
<223> Taqman UGT1A7 codon 139 E139-Fam (Forward)

<400> 26
atacttaaag gagagttgtt t 21

<210> 27
<211> 21
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(21)
<223> Taqman UGT1A7 codon 139 D139-Vic (Forward)

<400> 27
atacttaaag gacagttgtt t 21

<210> 28
<211> 31
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(31)
<223> Forward C3Y UGT1A9

<400> 28
gttctctgat ggcttacaca gggtggacca g 31

<210> 29
<211> 31
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(31)
<223> Reverse C3Y UGT1A9

<400> 29
ctggtccacc ctgtgtaagc catcagagaa c 31

<210> 30

<211> 34
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(34)
<223> Forward M33T UGT1A9

<400> 30
gctactggta gtgcccacgg atgggagcca ctgg 34

<210> 31
<211> 34
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(34)
<223> Reverse M33T UGT1A9

<400> 31
ccagtggctc ccatccgtgg gcactaccag tagc 34

<210> 32
<211> 45
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(45)
<223> Forward E139D UGT1A7

<400> 32
aattagtaga atacttaaag gacagttgtt ttgatgcagt gtttc 45

<210> 33
<211> 45
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(45)
<223> Reverse E139D UGT1A7

<400> 33
gaaacactgc atcaaaacaa ctgtccttta agtattctac taatt 45

<210> 34
<211> 23
<212> DNA
<213> Homo sapiens

<220>

<221> primer_bind
<222> (1)...(23)
<223> Forward G115S UGT1A7

<400> 34
gttcatccaa tagtattttt gac 23

<210> 35
<211> 23
<212> DNA
<213> Homo sapiens

<220>
<221> primer_bind
<222> (1)...(23)
<223> Reverse G115S UGT1A7

<400> 35
gtcaaaaata ctattggatg aac 23

<210> 36
<211> 2585
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2585)
<223> UGT1A9*1

<400> 36
atggcttgca cagggtggac cagccccctt cctctatgtg tgtgtctgct gctgacctgt 60
ggctttgccg aggcagggaa gctactggta gtgcccatgg atgggagcca ctggttcacc 120
atgaggtcgg tggtggagaa actcattctc agggggcatg aggtggttgt agtcatgcca 180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ttcaacttca 240
tataccctgg aggatctgga ccgggagttc aaggcttttg cccatgctca atggaaagca 300
caagtacgaa gtatatattc tctattaatg ggttcataca atgacatttt tgacttattt 360
ttttcaaatt gcaggagttt gtttaagac aaaaaattag tagaatactt aaaggagagt 420
tcttttgatg cagtgtttct cgatcctttt gataactgtg gcttaattgt tgccaaatat 480
ttctccctcc cctccgtggt cttcgccagg ggaatacttt gccactatct tgaagaaggt 540
gcacagtgcc ctgctcctct ttcctatgtc cccagaattc tcttagggtt ctcagatgcc 600
atgactttca aggagagagt acggaaccac atcatgcact tggaggaaca tttattatgc 660
caccgttttt tcaaaaatgc cctagaaata gcctctgaaa ttctccaaac acctgttacg 720
gagtatgatc tctacagcca cacatcaatt tggttgttgc gaacggactt tgttttggac 780
tatcccaaac ccgtgatgcc caacatgatc ttcattggtg gtatcaactg ccatcaggga 840
aagccgttgc ctatggaatt tgaagcctac attaatgctt ctggagaaca tggaattgtg 900
gttttctctt tgggatcaat ggtctcagaa attccagaga agaaagctat ggcaattgct 960
gatgctttgg gcaaaatccc tcagacagtc ctgtggcggt acactggaac ccgaccatcg 1020
aatcttgcga acaacacgat acttgttaag tggctacccc aaaacgatct gcttggtcac 1080
ccgatgaccc gtgcctttat cacccatgct ggttcccatg gtgtttatga aagcatatgc 1140
aatggcgttc ccatggtgat gatgcccttg tttggtgatc agatggacaa tgcaaagcgc 1200
atggagacta agggagctgg agtgaccctg aatgttctgg aaatgacttc tgaagattta 1260
gaaaatgctc taaaagcagt catcaatgac aaaagttaca aggagaacat catgcgcctc 1320
tccagccttc acaaggaccg cccggtggag ccgctggacc tggccgtgtt ctgggtggag 1380
tttgtgatga ggcacaaggg cgcgccacac ctgcgccccg cagcccacga cctcacctgg 1440
taccagtacc attccttgga cgtgattggt ttcctcttgg ccgtcgtgct gacagtggcc 1500
ttcatcacct ttaaatgttg tgcttatggc taccggaaat gcttggggaa aaaagggcga 1560

```
gttaagaaag cccacaaatc caagacccat tgagaagtgg gtgggaaata aggtaaaatt    1620
ttgaaccatt ccctagtcat ttccaaactt gaaaacagaa tcagtgttaa attcatttta    1680
ttcttattaa ggaaatactt tgcataaatt aatcagcccc agagtgcttt aaaaaattct    1740
cttaaataaa aataatagac tcgctagtca gtaaagatat ttgaatatgt atcgtgcccc    1800
ctctggtgtc tttgatcagg atgacatgtg ccatttttca gaggacgtgc agacaggctg    1860
gcattctaga ttacttttct tactctgaaa catggcctgt ttgggagtgc gggattcaaa    1920
ggtggtccca cggctgcccc tactgcaaat ggcagtttta atcttatctt ttggcttctg    1980
cagatggttg caattgatcc ttaaccaata atggtcagtc ctcatctctg tcgtgcttca    2040
taggtgccac cttgtgtgtt taaagaaggg aagctttgta cctttagagt gtaggtgaaa    2100
tgaatgaatg gcttggagtg cactgagaac agcatatgat ttcttgcttt ggggaaaaag    2160
aatgatgcta tgaaattggt gggtggtgta tttgagaaga taatcattgc ttatgtcaaa    2220
tggagctgaa tttgataaaa acccaaaata cagctatgaa gtgctgggca agtttacttt    2280
ttttctgatg tttcctacaa ctaaaaataa attaataaat ttatataaat tctatttaag    2340
tgttttcact ggtgtcgcat ttatttcttg ttaagttgca ttttctaatt acaaaagtaa    2400
tgcatgatta tgacagaaag tttggaaaat atagaggttc acacacacac gccttcattg    2460
cgtgtgcatg cataaatgca tgagaaaaga aaaataacca gtaatcacat cgcccagaaa    2520
taaccccagt tacaattgtg gcaaatacac atacttataa atattgcaga tatattaagt    2580
atacc                                                                2585
```

<210> 37
<211> 2585
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2585)
<223> UGT1A9*2

<400> 37

```
atggcttaca cagggtggac cagcccccctt cctctatgtg tgtgtctgct gctgacctgt      60
ggctttgccg aggcagggaa gctactggta gtgcccatgg atgggagcca ctggttcacc     120
atgaggtcgg tggtggagaa actcattctc agggggcatg aggtggttgt agtcatgcca     180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ttcaacttca     240
tataccctgg aggatctgga ccgggagttc aaggcttttg cccatgctca atggaaagca     300
caagtacgaa gtatatattc tctattaatg ggttcataca atgacatttt tgacttattt     360
ttttcaaatt gcaggagttt gtttaaagac aaaaaattag tagaatactt aaaggagagt     420
tcttttgatg cagtgtttct cgatcctttt gataactgtg gcttaattgt tgccaaatat     480
ttctccctcc cctccgtggt cttcgccagg ggaatacttt gccactatct tgaagaaggt     540
gcacagtgcc ctgctcctct ttcctatgtc cccagaattc tcttagggtt ctcagatgcc     600
atgactttca aggagagagt acggaaccac atcatgcact tggaggaaca tttattatgc     660
caccgttttt tcaaaaatgc cctagaaata gcctctgaaa ttctccaaac acctgttacg     720
gagtatgatc tctacagcca cacatcaatt tggttgttgc gaacggactt tgttttggac     780
tatcccaaac ccgtgatgcc caacatgatc ttcattggtg gtatcaactg ccatcaggga     840
aagccgttgc ctatggaatt tgaagcctac attaatgctt ctggagaaca tggaattgtg     900
gttttctctt tgggatcaat ggtctcagaa attccagaga agaaagctat ggcaattgct     960
gatgctttgg gcaaaatccc tcagacagtc ctgtggcggt acactggaac ccgaccatcg    1020
aatcttgcga acaacacgat acttgttaag tggctacccc aaaacgatct gcttggtcac    1080
ccgatgaccc gtgcctttat cacccatgct ggttcccatg gtgtttatga aagcatatgc    1140
aatggcgttc ccatggtgat gatgcccttg tttggtgatc agatggacaa tgcaaagcgc    1200
atggagacta agggagctgg agtgaccctg aatgttctgg aaatgacttc tgaagattta    1260
gaaaatgctc taaaagcagt catcaatgac aaaagttaca aggagaacat catgcgcctc    1320
tccagccttc acaaggaccg cccggtggag ccgctggacc tggccgtgtt ctgggtggag    1380
tttgtgatga ggcacaaggg cgcgccacac ctgcgccccg cagcccacga cctcacctgg    1440
taccagtacc attccttgga cgtgattggt ttcctcttgg ccgtcgtgct gacagtggcc    1500
ttcatcacct ttaaatgttg tgcttatggc taccggaaat gcttggggaa aaaagggcga    1560
gttaagaaag cccacaaatc caagacccat tgagaagtgg gtgggaaata aggtaaaatt    1620
```

```
ttgaaccatt ccctagtcat ttccaaactt gaaaacagaa tcagtgttaa attcatttta    1680
ttcttattaa ggaaatactt tgcataaatt aatcagcccc agagtgcttt aaaaaattct    1740
cttaaataaa aataatagac tcgctagtca gtaaagatat ttgaatatgt atcgtgcccc    1800
ctctggtgtc tttgatcagg atgacatgtg ccatttttca gaggacgtgc agacaggctg    1860
gcattctaga ttacttttct tactctgaaa catggcctgt ttgggagtgc gggattcaaa    1920
ggtggtccca cggctgcccc tactgcaaat ggcagtttta atcttatctt ttggcttctg    1980
cagatggttg caattgatcc ttaaccaata atggtcagtc ctcatctctg tcgtgcttca    2040
taggtgccac cttgtgtgtt taaagaaggg aagctttgta cctttagagt gtaggtgaaa    2100
tgaatgaatg gcttggagtg cactgagaac agcatatgat ttcttgcttt ggggaaaaag    2160
aatgatgcta tgaaattggt gggtggtgta tttgagaaga taatcattgc ttatgtcaaa    2220
tggagctgaa tttgataaaa acccaaaata cagctatgaa gtgctgggca agtttacttt    2280
ttttctgatg tttcctacaa ctaaaaataa attaataaat ttatataaat tctatttaag    2340
tgttttcact ggtgtcgcat ttatttcttg ttaagttgca ttttctaatt acaaaagtaa    2400
tgcatgatta tgacagaaag tttggaaaat atagaggttc acacacacac gccttcattg    2460
cgtgtgcatg cataaatgca tgagaaaaga aaaataacca gtaatcacat cgcccagaaa    2520
taaccccagt tacaattgtg gcaaatacac atacttataa atattgcaga tatattaagt    2580
atacc                                                                2585
```

<210> 38
<211> 2585
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2585)
<223> UGT1A9*3

<400> 38

```
atggcttgca cagggtggac cagcccccctt cctctatgtg tgtgtctgct gctgacctgt     60
ggctttgccg aggcagggaa gctactggta gtgcccacgg atgggagcca ctggttcacc    120
atgaggtcgg tggtggagaa actcattctc agggggcatg aggtggttgt agtcatgcca    180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ttcaacttca    240
tataccctgg aggatctgga ccgggagttc aaggcttttg cccatgctca atggaaagca    300
caagtacgaa gtatatattc tctattaatg ggttcataca atgacatttt tgacttattt    360
ttttcaaatt gcaggagttt gtttaaagac aaaaaattag tagaatactt aaaggagagt    420
tctttttgatg cagtgtttct cgatcctttt gataactgtg gcttaattgt tgccaaatat    480
ttctccctcc cctccgtggt cttcgccagg ggaatacttt gccactatct tgaagaaggt    540
gcacagtgcc ctgctcctct ttcctatgtc cccagaattc tcttagggtt ctcagatgcc    600
atgactttca aggagagagt acggaaccac atcatgcact tggaggaaca tttattatgc    660
caccgttttt tcaaaaatgc cctagaaata gcctctgaaa ttctccaaac acctgttacg    720
gagtatgatc tctacagcca cacatcaatt tggttgttgc gaacggactt tgttttggac    780
tatcccaaac ccgtgatgcc caacatgatc ttcattggtg gtatcaactg ccatcaggga    840
aagccgttgc ctatggaatt tgaagcctac attaatgctt ctggagaaca tggaattgtg    900
gttttctctt tgggatcaat ggtctcagaa attccagaga agaaagctat ggcaattgct    960
gatgctttgg gcaaaatccc tcagacagtc ctgtggcggt acactggaac ccgaccatcg   1020
aatcttgcga acaacacgat acttgttaag tggctacccc aaaacgatct gcttggtcac   1080
ccgatgaccc gtgcctttat cacccatgct ggttcccatg gtgtttatga aagcatatgc   1140
aatggcgttc ccatggtgat gatgcccttg tttggtgatc agatggacaa tgcaaagcgc   1200
atggagacta agggagctgg agtgaccctg aatgttctgg aaatgacttc tgaagattta   1260
gaaaatgctc taaaagcagt catcaatgac aaaagttaca aggagaacat catgcgcctc   1320
tccagccttc acaaggaccg cccggtggag ccgctggacc tggccgtgtt ctgggtggag   1380
tttgtgatga ggcacaaggg cgcgccacac ctgcgccccg cagcccacga cctcacctgg   1440
taccagtacc attccttgga cgtgattggt ttcctcttgg ccgtcgtgct gacagtggcc   1500
ttcatcacct ttaaatgttg tgcttatggc taccggaaat gcttggggaa aaaagggcga   1560
gttaagaaag cccacaaatc caagacccat tgagaagtgg gtgggaaata aggtaaaatt   1620
ttgaaccatt ccctagtcat ttccaaactt gaaaacagaa tcagtgttaa attcatttta   1680
```

```
ttcttattaa ggaaatactt tgcataaatt aatcagcccc agagtgcttt aaaaaattct    1740
cttaaataaa aataatagac tcgctagtca gtaaagatat ttgaatatgt atcgtgcccc    1800
ctctggtgtc tttgatcagg atgacatgtg ccatttttca gaggacgtgc agacaggctg    1860
gcattctaga ttacttttct tactctgaaa catggcctgt ttgggagtgc gggattcaaa    1920
ggtggtccca cggctgcccc tactgcaaat ggcagtttta atcttatctt ttggcttctg    1980
cagatggttg caattgatcc ttaaccaata atggtcagtc ctcatctctg tcgtgcttca    2040
taggtgccac cttgtgtgtt taaagaaggg aagctttgta cctttagagt gtaggtgaaa    2100
tgaatgaatg gcttggagtg cactgagaac agcatatgat ttcttgcttt ggggaaaaag    2160
aatgatgcta tgaaattggt gggtggtgta tttgagaaga taatcattgc ttatgtcaaa    2220
tggagctgaa tttgataaaa acccaaaata cagctatgaa gtgctgggca agtttacttt    2280
ttttctgatg tttcctacaa ctaaaaataa attaataaat ttatataaat tctatttaag    2340
tgttttcact ggtgtcgcat ttatttcttg ttaagttgca ttttctaatt acaaaagtaa    2400
tgcatgatta tgacagaaag tttggaaaat atagaggttc acacacacac gccttcattg    2460
cgtgtgcatg cataaatgca tgagaaaaga aaaataacca gtaatcacat cgcccagaaa    2520
taaccccagt tacaattgtg gcaaatacac atacttataa atattgcaga tatattaagt    2580
atacc                                                                2585

<210> 39
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 1

<400> 39
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
ttttttgcct ttgcaattct ttgaaaatgc tttttttacag tactagtcct tcttccccat    720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccacctttttg ttgctgaatt agagatattg ggtttgcagg   1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atctttttccc tttaaggctt ggaggctagc   1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg    1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
```

```
gaggtcagtg ctaagggcct tgttttcttt gcttagagta tgagttgcca tcttctctgg   1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa   1860
aaaaaattag ctttaatcaa atttactctt actttatctt tctgaacctt caaggtccaa   1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc   1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca   2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gattttttt    2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca    2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat   2220
ggcttgcaca gggtggacca gccccttcc tctatgtgtg tgtctgctgc tgacctgtgg    2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat   2340
gaggtcggtg gtggagaaac tcattctcag gg                                 2372
```

<210> 40
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 2

<400> 40

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc     60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc    120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag    180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg    240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt    300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca    360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa    420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag    480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga    540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag    600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac    660
ttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat    720
ttgcttttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta    780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc    840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc    900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat    960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg   1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca   1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac   1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt   1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc   1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg   1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc   1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg   1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag   1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga   1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag   1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa   1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag   1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagta tgagttgcca tcttctctgg   1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa   1860
aaaaaattag ctttaatcaa atttactctt actttatctt tctgaacctt caaggtccaa   1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc   1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca   2040
```

```
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt     2100
tatgaaagga taaaaacacg ccctctattg gagtcaggtt ttgtgctggt atttctccca     2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat     2220
ggcttgcaca gggtggacca gccccttcc tctatgtgtg tgtctgctgc tgacctgtgg     2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat     2340
gaggtcggtg gtggagaaac tcattctcag gg                                   2372

&lt;210&gt; 41
&lt;211&gt; 2372
&lt;212&gt; DNA
&lt;213&gt; Homo sapiens

&lt;220&gt;
&lt;221&gt; allele
&lt;222&gt; (1)...(2372)
&lt;223&gt; UGT1A9 Haplotype 3

&lt;400&gt; 41
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc       60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc      120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag      180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg      240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt      300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca      360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa      420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag      480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga      540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag      600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac      660
ttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat      720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta      780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc      840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc      900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat      960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg     1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca     1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac     1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt     1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc     1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg     1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc     1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg     1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag     1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga     1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag     1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa     1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag     1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg     1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa     1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa     1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc     1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca     2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt     2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca      2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat     2220
ggcttgcaca gggtggacca gccccttcc tctatgtgtg tgtctgctgc tgacctgtgg     2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat     2340
```

gaggtcggtg gtggagaaac tcattctcag gg 2372

<210> 42
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 4

<400> 42
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc 60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc 120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag 180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg 240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt 300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca 360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa 420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag 480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga 540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag 600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac 660
tttttttgcct ttgcaattct ttgaaaatgc tttttacag tactagtcct tcttccccat 720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta 780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc 840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc 900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat 960
atcttgaaag accatatccc ccacctttg ttgctgaatt agagatattg ggtttgcagg 1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca 1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac 1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt 1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc 1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg 1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc 1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catatttcc tgaaggaggg 1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag 1500
cagagtgctc tcgcaaggat tgggcgggca acttccccact gcgtgcgatg tattttagga 1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag 1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa 1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag 1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg 1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa 1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa 1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc 1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca 2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gattttttt 2100
tatgaaagga taaaaacacg ccctctattg gggtcaggtt ttgtgctggt atttctccca 2160
cctactgtat catagagct tagattccca gctgcttgct ctcagctgca gttctctgat 2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg 2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat 2340
gaggtcggtg gtggagaaac tcattctcag gg 2372

<210> 43
<211> 2372
<212> DNA

<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UHT1A9 Haplotype 5

<400> 43
```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca caatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
tttttttgcct ttgcaattct ttgaaaatgc tttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccacctttttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg    1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tattttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gattttttttt    2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca    2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat    2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg    2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                   2372
```

<210> 44
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)

<223> UGT1A9 Haplotype 6

<400> 44
```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cgtaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
tttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat    720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctagggggca   1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catatttcc tgaaggaggg     1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt    2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca     2160
cctactgtat catagagct tagattccca gctgcttgct ctcagctgca gttctctgat     2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg   2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                  2372
```

<210> 45
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 7

<400> 45
```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
```

accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag 180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg 240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt 300
actattaata gcctactgtg cactagaagc cgtaccaata acagaaacag ttgcttaaca 360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa 420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag 480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga 540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag 600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac 660
ttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat 720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta 780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc 840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc 900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat 960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg 1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca 1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac 1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt 1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc 1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg 1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc 1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catatttcc tgaaggaggg 1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag 1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tattttagga 1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag 1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa 1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag 1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg 1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa 1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa 1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc 1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca 2040
aaagctactc atatattctt gttctttttgg gtaaatcatt gtcagtgact gatttttttt 2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca 2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat 2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg 2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat 2340
gaggtcggtg gtggagaaac tcattctcag gg 2372

<210> 46
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 8

<400> 46
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc 60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc 120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag 180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg 240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt 300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca 360
catatttggc atgttatatg tgttatatac tgtattatca caatgaagtc agctagagaa 420

```
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
tttttgcct ttgcaattct ttgaaaatgc tttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg    1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgtttcttt gcttagagca tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt    2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca    2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat    2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg    2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                   2372
```

<210> 47
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 9

<400> 47

```
ctgttttgcc tgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
tttttgcct ttgcaattct ttgaaaatgc tttttacag tactagtcct tcttccccat     720
```

```
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta    780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc    840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc    900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat    960
atcttgaaag accatatccc ccacctttttg ttgctgaatt agagatattg ggtttgcagg  1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca   1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac   1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt   1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc   1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg   1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc   1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg   1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag   1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tattttagga   1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag   1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa   1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag   1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg   1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa   1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa   1920
aagcattggt taataattct gctactaaac ttaacattgc agcacagggc atgttctgcc   1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca   2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt   2100
tatgaaagga taaaaacacg ccctctattg gggtcaggtt ttgtgctggt atttctccca   2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat   2220
ggcttgcaca gggtggacca gccccttcc tctatgtgtg tgtctgctgc tgacctgtgg    2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat   2340
gaggtcggtg gtggagaaac tcattctcag gg                                  2372
```

<210> 48
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 10

<400> 48

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc     60
cgggttcaag tgattcttct gcctcagcct ccagagtagc tgggattaca ggcatgcacc    120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag    180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg    240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt    300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca    360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa    420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag    480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga    540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag    600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac    660
tttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta    780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc    840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc    900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat    960
atcttgaaag accatatccc ccacctttttg ttgctgaatt agagatattg ggtttgcagg  1020
```

caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca 1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac 1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt 1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc 1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg 1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc 1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg 1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag 1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tattttagga 1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag 1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa 1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag 1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg 1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa 1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa 1920
aagcattggt taataattct gctactaaac ttaacattgc agcacagggc atgttctgcc 1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca 2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gattttttt 2100
tatgaaagga taaaaacacg ccctctattg gggtcaggtt ttgtgctggt atttctccca 2160
cctactgtat catagagagct tagattccca gctgcttgct ctcagctgca gttctctgat 2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg 2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat 2340
gaggtcggtg gtggagaaac tcattctcag gg 2372

<210> 49
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 11

<400> 49
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc 60
cgggtttaag tgattctcct gcctcagcct cccagagtagc tgggattaca ggcatgcacc 120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag 180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg 240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt 300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca 360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa 420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag 480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga 540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag 600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac 660
ttttttgcct ttgcaattct ttgaaaatgc tttttacag tactagtcct tcttccccat 720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta 780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc 840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc 900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat 960
atcttgaaag accatatccc ccacctttg ttgctgaatt agagatattg ggtttgcagg 1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca 1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac 1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt 1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc 1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg 1320

```
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc 1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg 1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag 1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tattttagga 1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag 1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa 1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag 1740
gaggtcagtg ctaagggcct tgtttctttt gcttagagca tgagttgcca tcttctctgg 1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa 1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa 1920
aagcattggt taataattct gctactaaac ttaacattgc agcacagggc atgttctgcc 1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca 2040
aaagctactc atatattctt gttctttggg gtaaatcatt gtcagtgact gatttttttt 2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca 2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat 2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg 2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat 2340
gaggtcggtg gtggagaaac tcattctcag gg 2372
```

<210> 50
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 12

<400> 50

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc 60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc 120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag 180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg 240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt 300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca 360
catatttggc atgttatatg tgttatatac tgtattatca caatgaagtc agctagagaa 420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag 480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga 540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag 600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac 660
ttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat 720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta 780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc 840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc 900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat 960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg 1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca 1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac 1140
ttcaggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt 1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc 1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg 1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc 1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg 1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag 1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga 1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag 1620
```

```
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt    2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca     2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat    2220
ggcttgcaca gggtggacca gccccttcc tctatgtgtg tgtctgctgc tgacctgtgg     2280
ctttgccgag gcagggaagc tactggtagt gcccacggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                  2372
```

<210> 51
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 13

<400> 51

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cgtaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
ttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg    1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagta tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa    1920
```

```
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gattttttt     2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca     2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat    2220
ggcttgcaca gggtggacca gccccttcc tctatgtgtg tgtctgctgc tgacctgtgg     2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                  2372
```

<210> 52
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2)
<223> UGT1A9 Haplotype 14

<400> 52

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca caatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
tttttgcct ttgcaattct ttgaaaatgc tttttacag tactagtcct tcttccccat       720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catatttcc tgaaggaggg     1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgtttcttt gcttagagta tgagttgcca tcttctctgg     1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactctt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gattttttt     2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca     2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat    2220
```

```
ggcttgcaca gggtggacca gccccсttcc tctatgtgtg tgtctgctgc tgacctgtgg    2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                  2372
```

<210> 53
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 15

<400> 53

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattcttct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cgtaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
tttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg    1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tattttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gctactaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt    2100
tatgaaagga taaaaacacg ccctctattg gggtcaggtt ttgtgctggt atttctccca    2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat    2220
ggcttgcaca gggtggacca gcccccttcc tctatgtgtg tgtctgctgc tgacctgtgg    2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                  2372
```

<210> 54

<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 16

<400> 54

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattcttct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cgtaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
tttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg    1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tattttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gattttttt     2100
tatgaaagga taaaaacacg ccctctattg gggtcaggtt ttgtgctggt atttctccca    2160
cctactgtat catagagact tagattccca gctgcttgct ctcagctgca gttctctgat    2220
ggcttgcaca gggtggacca gccccttcc tctatgtgtg tgtctgctgc tgacctgtgg     2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                 2372
```

<210> 55
<211> 2372
<212> DNA
<213> Homo sapiens

<220>

<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 17

<400> 55
```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
tttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catatttccc tgaaggaggg    1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tattttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gcccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgtttctttt gcttagagca tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gctactaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt    2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca     2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat    2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg    2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                 2372
```

<210> 56
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 18

<400> 56

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc     60
cgggtttaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc    120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag    180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg    240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt    300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca    360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa    420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag    480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga    540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag    600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac    660
tttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat    720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta    780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc    840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc    900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat    960
atcttgaaag accatatccc ccacctttttg ttgctgaatt agagatattg ggtttgcagg   1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca   1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac   1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt   1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc   1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg   1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc   1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg   1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag   1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga   1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag   1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa   1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag   1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg   1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa   1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa   1920
aagcattggt taataattct gctactaaac ttaacattgc agcacagggc atgttctgcc   1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca   2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt   2100
tatgaaagga taaaaacacg ccctctattg gggtcaggtt ttgtgctggt atttctccca   2160
cctactgtat catagagct tagattccca gctgcttgct ctcagctgca gttctctgat   2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg  2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat   2340
gaggtcggtg gtggagaaac tcattctcag gg                                   2372
```

<210> 57
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haploytpe 19

<400> 57

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc     60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc    120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag    180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg    240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt    300
```

```
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca    360
catatttggc atgttatatg tgttatatac tgtattatca taatgaagtc agctagagaa    420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag    480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga    540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag    600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac    660
tttttgcct ttgcaattct ttgaaaatgc tttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta    780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc    840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc    900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat    960
atcttgaaag accatatccc ccacctttg ttgctgaatt agagatattg ggtttgcagg     1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca   1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac   1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt   1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc   1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg   1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc   1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg   1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag   1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga   1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag   1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa   1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag   1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg   1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa   1860
aaaaaattag ctttaatcaa atttactctt actttatctt tctgaacctt caaggtccaa   1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc   1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca   2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt   2100
tatgaaagga taaaaacacg ccctctattg ggtcaggttt ttgtgctggt atttctccca   2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat   2220
ggcttgcaca gggtggacca gcccccttcc tctatgtgtg tgtctgctgc tgacctgtgg   2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat   2340
gaggtcggtg gtggagaaac tcattctcag gg                                  2372
```

<210> 58
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 20

<400> 58

```
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc     60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc    120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag    180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg    240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt    300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca    360
catatttggc atgttatatg tgttatatac tgtattatca caatgaagtc agctagagaa    420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag    480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga    540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag    600
```

```
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
ttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat     960
atcttgaaag accatatccc ccaccttttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catatttccc tgaaggaggg    1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tatcttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactctt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt    2100
tatgaaagga taaaaacacg ccctctattg gggtcaggtt ttgtgctggt atttctccca    2160
cctactgtat cataggagct tagattccca gctgcttgct ctcagctgca gttctctgat    2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg    2280
ctttgccgag gcagggaagc tactggtagt gcccatggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                  2372

<210> 59
<211> 2372
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(2372)
<223> UGT1A9 Haplotype 21

<400> 59
ctgttttgcc cgggctggag tataatggcg tgatctcagc tcaatgcaac ctccgcttcc      60
cgggttcaag tgattctcct gcctcagcct ccagagtagc tgggattaca ggcatgcacc     120
accacctgca gctaattttt tgcattttta gtagagatag ggtttcacca tgttggccag     180
gctggtctcc aactcctggc ctcccgtgat acgcccacct tgacctccca aagtgctggg     240
actacaggtg tgagccacca cgcccaggca cacatagaat ttttgactcc ctaaaaattt     300
actattaata gcctactgtg cactagaagc cttaccaata acagaaacag ttgcttaaca     360
catatttggc atgttatatg tgttatatac tgtattatca caatgaagtc agctagagaa     420
aagaaaatgt tattaagaaa atcttaagga agagaaaatt aagtattcat taagtggaag     480
tggatcatga taaaggtctt cctcttgatt gtcctccatt gagtaggctg agaaggagga     540
agaggtgggt tggttttgct gtttcagggg tggcagaggg ggaagaagtg gaggaagaag     600
gaggagagac aggtacactt ggtgtaactt tacagaatta catcataatt attatttgac     660
ttttttgcct ttgcaattct ttgaaaatgc ttttttacag tactagtcct tcttccccat     720
ttgctttagt ttcagtgccc attcatggaa gggtttgtgt tgtaaaataa gtcaaaagta     780
gtcttaataa ttggaagcct ttgccaaact gtttaatagg aatttgtttt ctggcatggc     840
ttcttctatg tcttctttag tatctggtac tgattcaaaa gcactcatct ccatcaagtc     900
```

```
atcttctgtt gattcctctg gtgtggtgtc tattcattct tgaatttctc acagattcat    960
atcttgaaag accatatccc ccacctttg ttgctgaatt agagatattg ggtttgcagg    1020
caagtagacc actttgacac cttcagtgtt gaactcatgg gttctgggtg gctaggggca    1080
ttgtccaaaa atcaaaagaa ctttgaaaga ccgtctctta ctggcaagat attacctgac    1140
ttcagggaca aagtaatgat agaaccaatc cagaaaaagt gttcttgccg aggccttctt    1200
gtacaacaaa aaaactggca gtgggtattg atcttttccc tttaaggctt ggaggctagc    1260
aggcttatgg atgggggcag tcctatttgt aaacccaaac atacaaacat acaaactatg    1320
tcaaaggcat agcatgggta ctgtgaaagg agggtgaaaa cacaaagttg acatcacctc    1380
tgacctcaag gagtgctcag cagactgaga gagacaagta catattttcc tgaaggaggg    1440
cactggagtg atggcgtgtt tagaatgtgc aagttgagcg gtcactgaga ggcagctcag    1500
cagagtgctc tcgcaaggat tgggcgggca acttcccact gcgtgcgatg tattttagga    1560
aagccattta aaataggaga cggttacttt ccatcaagtc cctggtatgg tccatggaag    1620
cagggttgtc agtctcattt cagcatttta gaggcttctc agggtttgga aatggaagaa    1680
gagaagcagc aatatgtatg cattgcagag acacaggcga gccccaattt aggaggttag    1740
gaggtcagtg ctaagggcct tgttttcttt gcttagagca tgagttgcca tcttctctgg    1800
acagagagta tttggttgcc taaaggtaaa atctaaattt tgctctggga caaattccaa    1860
aaaaaattag ctttaatcaa atttactttt actttatctt tctgaacctt caaggtccaa    1920
aagcattggt taataattct gcttctaaac ttaacattgc agcacagggc atgttctgcc    1980
cccaaggcaa agaccataag ctactgttgt ctggaaaaca tacaaataga tatctcagca    2040
aaagctactc atatattctt gttcttttgg gtaaatcatt gtcagtgact gatttttttt    2100
tatgaaagga taaaaacacg ccctctattg ggtcaggtt ttgtgctggt atttctccca    2160
cctactgtat catagggagct tagattccca gctgcttgct ctcagctgca gttctctgat    2220
ggcttgcaca gggtggacca gccccccttcc tctatgtgtg tgtctgctgc tgacctgtgg    2280
ctttgccgag gcagggaagc tactggtagt gcccacggat gggagccact ggttcaccat    2340
gaggtcggtg gtggagaaac tcattctcag gg                                  2372
```

<210> 60
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(1229)
<223> UGT1A7*1

<400> 60

```
atggctcgtg cagggtggac tggcctcctt ccactatatg tgtgtctact gctgacctgt      60
ggctttgcca aggcagggaa gctgctggta gtgcccatgg atgggagcca ctggttcacc     120
atgcagtcgg tggtggagaa actcatcctc agggggcatg aggtggtcgt agtcatgcca     180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ctcaacctca     240
tacactctgg aggatcagga ccgggagttc atggttttg ccgatgctcg ctggacggca      300
ccattgcgaa gtgcattttc tctattaaca agttcatcca atggtatttt tgacttattt     360
ttttcaaatt gcaggagttt gtttaatgac cgaaaattag tagaatactt aaaggagagt     420
tgttttgatg cagtgtttct cgatcctttt gatgcctgtg gcttaattgt tgccaaatat     480
ttctccctcc cctctgtggt cttcgccagg ggaatatttt gccactatct tgaagaaggt     540
gcacagtgcc ctgctcctct ttcctatgtc cccagacttc tcttagggtt ctcagacgcc     600
atgactttca aggagagagt atggaaccac atcatgcact tggaggaaca tttattttgc     660
ccctattttt tcaaaaatgt cttagaaata gcctctgaaa ttctccaaac ccctgtcacg     720
gcatatgatc tctacagcca cacatcaatt tggttgttgc gaactgactt tgttttggag     780
tatcccaaac ccgtgatgcc caatatgatc ttcattggtg gtatcaactg tcatcaggga     840
aagccagtgc ctatggtaag ttatctcccc tttagcacat taagaataat ctggctttgg     900
aaattaaaag atttcttaca gaatcataat ttatcattta catttgtccc atttggaatt     960
tctttctggt ttaaggaatt cttttgtacc aattcactta attgttgggt agcaaattgt    1020
ataaagcagc tcttgttgat atgtaagtgt atacaattga tataattgta gatcatatct    1080
aggctgcaat ctaaatgcta tttttggaaa aatacaaaaa aaccacagta agaaatgaaa    1140
cttccctttt tttgctaatt ctacactacc cccagaggaa aatattctta gcagttttgt    1200
```

gtgaattgtt ttcaattttt ttgaaatta 1229

<210> 61
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(1229)
<223> UGT1A7*2

<400> 61
atggctcgtg cagggtggac tggcctcctt ccactatatg tgtgtctact gctgacctgt 60
ggctttgcca aggcagggaa gctgctggta gtgcccatgg atgggagcca ctggttcacc 120
atgcagtcgg tggtggagaa actcatcctc agggggcatg aggtggtcgt agtcatgcca 180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ctcaacctca 240
tacactctgg aggatcagga ccgggagttc atggtttttg ccgatgctcg ctggacggca 300
ccattgcgaa gtgcattttc tctattaaca agttcatcca atggtatttt tgacttattt 360
ttttcaaatt gcaggagttt gtttaaggac aaaaaattag tagaatactt aaaggagagt 420
tgttttgatg cagtgtttct cgatcctttt gatgcctgtg gcttaattgt tgccaaatat 480
ttctccctcc cctctgtggt cttcgccagg ggaatatttt gccactatct tgaagaaggt 540
gcacagtgcc ctgctcctct ttcctatgtc cccagacttc tcttagggtt ctcagacgcc 600
atgactttca aggagagagt atggaaccac atcatgcact tggaggaaca tttattttgc 660
ccctattttt tcaaaaatgt cttagaaata gcctctgaaa ttctccaaac ccctgtcacg 720
gcatatgatc tctacagcca cacatcaatt tggttgttgc gaactgactt tgttttggag 780
tatcccaaac ccgtgatgcc caatatgatc ttcattggtg gtatcaactg tcatcaggga 840
aagccagtgc ctatggtaag ttatctcccc tttagcacat taagaataat ctggctttgg 900
aaattaaaag atttcttaca gaatcataat ttatcattta catttgtccc atttggaatt 960
tctttctggt ttaaggaatt cttttgtacc aattcactta attgttgggt agcaaattgt 1020
ataaagcagc tcttgttgat atgtaagtgt atacaattga tataattgta gatcatatct 1080
aggctgcaat ctaaatgcta tttttggaaa aatacaaaaa aaccacagta agaaatgaaa 1140
cttccctttt tttgctaatt ctacactacc cccagaggaa aatattctta gcagttttgt 1200
gtgaattgtt ttcaattttt ttgaaatta 1229

<210> 62
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(1229)
<223> UGT1A7*3

<400> 62
atggctcgtg cagggtggac tggcctcctt ccactatatg tgtgtctact gctgacctgt 60
ggctttgcca aggcagggaa gctgctggta gtgcccatgg atgggagcca ctggttcacc 120
atgcagtcgg tggtggagaa actcatcctc agggggcatg aggtggtcgt agtcatgcca 180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ctcaacctca 240
tacactctgg aggatcagga ccgggagttc atggtttttg ccgatgctcg ctggacggca 300
ccattgcgaa gtgcattttc tctattaaca agttcatcca atggtatttt tgacttattt 360
ttttcaaatt gcaggagttt gtttaaggac aaaaaattag tagaatactt aaaggagagt 420
tgttttgatg cagtgtttct cgatcctttt gatgcctgtg gcttaattgt tgccaaatat 480
ttctccctcc cctctgtggt cttcgccagg ggaatatttt gccactatct tgaagaaggt 540
gcacagtgcc ctgctcctct ttcctatgtc cccagacttc tcttagggtt ctcagacgcc 600
atgactttca aggagagagt acggaaccac atcatgcact tggaggaaca tttattttgc 660

ccctattttt tcaaaaatgt cttagaaata gcctctgaaa ttctccaaac ccctgtcacg 720
gcatatgatc tctacagcca cacatcaatt tggttgttgc gaactgactt tgttttggag 780
tatcccaaac ccgtgatgcc caatatgatc ttcattggtg gtatcaactg tcatcaggga 840
aagccagtgc ctatggtaag ttatctcccc tttagcacat taagaataat ctggctttgg 900
aaattaaaag atttcttaca gaatcataat ttatcattta catttgtccc atttggaatt 960
tctttctggt ttaaggaatt cttttgtacc aattcactta attgttgggt agcaaattgt 1020
ataaagcagc tcttgttgat atgtaagtgt atacaattga tataattgta gatcatatct 1080
aggctgcaat ctaaatgcta tttttggaaa aatacaaaaa aaccacagta agaaatgaaa 1140
cttccctttt tttgctaatt ctacactacc cccagaggaa aatattctta gcagttttgt 1200
gtgaattgtt ttcaattttt ttgaaatta 1229

<210> 63
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(1229)
<223> UGT1A7*4

<400> 63
atggctcgtg cagggtggac tggcctcctt ccactatatg tgtgtctact gctgacctgt 60
ggctttgcca aggcagggaa gctgctggta gtgcccatgg atgggagcca ctggttcacc 120
atgcagtcgg tggtggagaa actcatcctc agggggcatg aggtggtcgt agtcatgcca 180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ctcaacctca 240
tacactctgg aggatcagga ccgggagttc atggtttttg ccgatgctcg ctggacggca 300
ccattgcgaa gtgcattttc tctattaaca agttcatcca atggtatttt tgacttattt 360
ttttcaaatt gcaggagttt gtttaatgac cgaaaattag tagaatactt aaaggagagt 420
tgttttgatg cagtgtttct cgatcctttt gatgcctgtg gcttaattgt tgccaaatat 480
ttctccctcc cctctgtggt cttcgccagg ggaatatttt gccactatct tgaagaaggt 540
gcacagtgcc ctgctcctct ttcctatgtc cccagacttc tcttagggtt ctcagacgcc 600
atgactttca aggagagagt acggaaccac atcatgcact tggaggaaca tttatttgc 660
ccctattttt tcaaaaatgt cttagaaata gcctctgaaa ttctccaaac ccctgtcacg 720
gcatatgatc tctacagcca cacatcaatt tggttgttgc gaactgactt tgttttggag 780
tatcccaaac ccgtgatgcc caatatgatc ttcattggtg gtatcaactg tcatcaggga 840
aagccagtgc ctatggtaag ttatctcccc tttagcacat taagaataat ctggctttgg 900
aaattaaaag atttcttaca gaatcataat ttatcattta catttgtccc atttggaatt 960
tctttctggt ttaaggaatt cttttgtacc aattcactta attgttgggt agcaaattgt 1020
ataaagcagc tcttgttgat atgtaagtgt atacaattga tataattgta gatcatatct 1080
aggctgcaat ctaaatgcta tttttggaaa aatacaaaaa aaccacagta agaaatgaaa 1140
cttccctttt tttgctaatt ctacactacc cccagaggaa aatattctta gcagttttgt 1200
gtgaattgtt ttcaattttt ttgaaatta 1229

<210> 64
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(1229)
<223> UGT1A7*5

<400> 64
atggctcgtg cagggtggac tggcctcctt ccactatatg tgtgtctact gctgacctgt 60
ggctttgcca aggcagggaa gctgctggta gtgcccatgg atgggagcca ctggttcacc 120

```
atgcagtcgg tggtggagaa actcatcctc agggggcatg aggtggtcgt agtcatgcca     180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ctcaacctca     240
tacactctgg aggatcagga ccgggagttc atggttttg ccgatgctcg ctggacggca      300
ccattgcgaa gtgcattttc tctattaaca agttcatcca atagtatttt tgacttattt     360
ttttcaaatt gcaggagttt gtttaatgac cgaaaattag tagaatactt aaaggagagt     420
tgttttgatg cagtgtttct cgatcctttt gatgcctgtg gcttaattgt tgccaaatat     480
ttctccctcc cctctgtggt cttcgccagg ggaatatttt gccactatct tgaagaaggt     540
gcacagtgcc ctgctcctct ttcctatgtc cccagacttc tcttagggtt ctcagacgcc     600
atgactttca aggagagagt atggaaccac atcatgcact tggaggaaca tttattttgc     660
ccctattttt tcaaaaatgt cttagaaata gcctctgaaa ttctccaaac ccctgtcacg     720
gcatatgatc tctacagcca cacatcaatt tggttgttgc gaactgactt tgttttggag     780
tatcccaaac ccgtgatgcc caatatgatc ttcattggtg gtatcaactg tcatcaggga     840
aagccagtgc ctatggtaag ttatctcccc tttagcacat taagaataat ctggctttgg     900
aaattaaaag atttcttaca gaatcataat ttatcattta catttgtccc atttggaatt     960
tctttctggt ttaaggaatt cttttgtacc aattcactta attgttgggt agcaaattgt    1020
ataaagcagc tcttgttgat atgtaagtgt atacaattga tataattgta gatcatatct    1080
aggctgcaat ctaaatgcta tttttggaaa aatacaaaaa aaccacagta agaaatgaaa    1140
cttccctttt tttgctaatt ctacactacc cccagaggaa aatattctta gcagttttgt    1200
gtgaattgtt ttcaattttt ttgaaatta                                     1229
```

<210> 65
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(1229)
<223> UGT1A7*6

<400> 65

```
atggctcgtg cagggtggac tggcctcctt ccactatatg tgtgtctact gctgacctgt      60
ggctttgcca aggcagggaa gctgctggta gtgcccatgg atgggagcca ctggttcacc     120
atgcagtcgg tggtggagaa actcatcctc agggggcatg aggtggtcgt agtcatgcca     180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ctcaacctca     240
tacactctgg aggatcagga ccgggagttc atggttttg ccgatgctcg ctggacggca      300
ccattgcgaa gtgcattttc tctattaaca agttcatcca atggtatttt tgacttattt     360
ttttcaaatt gcaggagttt gtttaatgac cgaaaattag tagaatactt aaaggacagt     420
tgttttgatg cagtgtttct cgatcctttt gatgcctgtg gcttaattgt tgccaaatat     480
ttctccctcc cctctgtggt cttcgccagg ggaatatttt gccactatct tgaagaaggt     540
gcacagtgcc ctgctcctct ttcctatgtc cccagacttc tcttagggtt ctcagacgcc     600
atgactttca aggagagagt atggaaccac atcatgcact tggaggaaca tttattttgc     660
ccctattttt tcaaaaatgt cttagaaata gcctctgaaa ttctccaaac ccctgtcacg     720
gcatatgatc tctacagcca cacatcaatt tggttgttgc gaactgactt tgttttggag     780
tatcccaaac ccgtgatgcc caatatgatc ttcattggtg gtatcaactg tcatcaggga     840
aagccagtgc ctatggtaag ttatctcccc tttagcacat taagaataat ctggctttgg     900
aaattaaaag atttcttaca gaatcataat ttatcattta catttgtccc atttggaatt     960
tctttctggt ttaaggaatt cttttgtacc aattcactta attgttgggt agcaaattgt    1020
ataaagcagc tcttgttgat atgtaagtgt atacaattga tataattgta gatcatatct    1080
aggctgcaat ctaaatgcta tttttggaaa aatacaaaaa aaccacagta agaaatgaaa    1140
cttccctttt tttgctaatt ctacactacc cccagaggaa aatattctta gcagttttgt    1200
gtgaattgtt ttcaattttt ttgaaatta                                     1229
```

<210> 66
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(1229)
<223> UGT1A7*7

<400> 66
atggctcgtg cagggtggac tggcctcctt ccactatatg tgtgtctact gctgacctgt 60
ggctttgcca aggcagggaa gctgctggta gtgcccatgg atgggagcca ctggttcacc 120
atgcagtcgg tggtggagaa actcatcctc agggggcatg aggtggtcgt agtcatgcca 180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ctcaacctca 240
tacactctgg aggatcagga ccgggagttc atggtttttg ccgatgctcg ctggacggca 300
ccattgcgaa gtgcattttc tctattaaca agttcatcca atggtatttt tgacttattt 360
ttttcaaatt gcaggagttt gtttaaggac aaaaaattag tagaatactt aaaggacagt 420
tgttttgatg cagtgtttct cgatcctttt gatgcctgtg gcttaattgt tgccaaatat 480
ttctccctcc cctctgtggt cttcgccagg ggaatatttt gccactatct tgaagaaggt 540
gcacagtgcc ctgctcctct ttcctatgtc cccagacttc tcttagggtt ctcagacgcc 600
atgactttca aggagagagt atggaaccac atcatgcact tggaggaaca tttattttgc 660
ccctattttt tcaaaaatgt cttagaaata gcctctgaaa ttctccaaac ccctgtcacg 720
gcatatgatc tctacagcca cacatcaatt tggttgttgc gaactgactt tgttttggag 780
tatcccaaac ccgtgatgcc caatatgatc ttcattggtg gtatcaactg tcatcaggga 840
aagccagtgc ctatggtaag ttatctcccc tttagcacat taagaataat ctggctttgg 900
aaattaaaag atttcttaca gaatcataat ttatcattta catttgtccc atttggaatt 960
tctttctggt ttaaggaatt cttttgtacc aattcactta attgttgggt agcaaattgt 1020
ataaagcagc tcttgttgat atgtaagtgt atacaattga tataattgta gatcatatct 1080
aggctgcaat ctaaatgcta tttttggaaa aatacaaaaa aaccacagta agaaatgaaa 1140
cttccctttt tttgctaatt ctacactacc cccagaggaa aatattctta gcagttttgt 1200
gtgaattgtt ttcaattttt ttgaaatta 1229

<210> 67
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(1229)
<223> UGT1A7*8

<400> 67
atggctcgtg cagggtggac tggcctcctt ccactatatg tgtgtctact gctgacctgt 60
ggctttgcca aggcagggaa gctgctggta gtgcccatgg atgggagcca ctggttcacc 120
atgcagtcgg tggtggagaa actcatcctc agggggcatg aggtggtcgt agtcatgcca 180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ctcaacctca 240
tacactctgg aggatcagga ccgggagttc atggtttttg ccgatgctcg ctggacggca 300
ccattgcgaa gtgcattttc tctattaaca agttcatcca atggtatttt tgacttattt 360
ttttcaaatt gcaggagttt gtttaaggac aaaaaattag tagaatactt aaaggacagt 420
tgttttgatg cagtgtttct cgatcctttt gatgcctgtg gcttaattgt tgccaaatat 480
ttctccctcc cctctgtggt cttcgccagg ggaatatttt gccactatct tgaagaaggt 540
gcacagtgcc ctgctcctct ttcctatgtc cccagacttc tcttagggtt ctcagacgcc 600
atgactttca aggagagagt acggaaccac atcatgcact tggaggaaca tttattttgc 660
ccctattttt tcaaaaatgt cttagaaata gcctctgaaa ttctccaaac ccctgtcacg 720
gcatatgatc tctacagcca cacatcaatt tggttgttgc gaactgactt tgttttggag 780
tatcccaaac ccgtgatgcc caatatgatc ttcattggtg gtatcaactg tcatcaggga 840
aagccagtgc ctatggtaag ttatctcccc tttagcacat taagaataat ctggctttgg 900
aaattaaaag atttcttaca gaatcataat ttatcattta catttgtccc atttggaatt 960
tctttctggt ttaaggaatt cttttgtacc aattcactta attgttgggt agcaaattgt 1020

```
ataaagcagc tcttgttgat atgtaagtgt atacaattga tataattgta gatcatatct    1080
aggctgcaat ctaaatgcta tttttggaaa aatacaaaaa aaccacagta agaaatgaaa    1140
cttccctttt tttgctaatt ctacactacc cccagaggaa aatattctta gcagttttgt    1200
gtgaattgtt ttcaattttt ttgaaatta                                      1229
```

<210> 68
<211> 1229
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (1)...(1229)
<223> UGT1A7*9

<400> 68

```
atggctcgtg cagggtggac tggcctcctt ccactatatg tgtgtctact gctgacctgt      60
ggctttgcca aggcagggaa gctgctggta gtgcccatgg atgggagcca ctggttcacc     120
atgcagtcgg tggtggagaa actcatcctc agggggcatg aggtggtcgt agtcatgcca     180
gaggtgagtt ggcaactggg aagatcactg aattgcacag tgaagactta ctcaacctca     240
tacactctgg aggatcagga ccgggagttc atggtttttg ccgatgctcg ctggacggca     300
ccattgcgaa gtgcattttc tctattaaca agttcatcca atagtatttt tgacttattt     360
ttttcaaatt gcaggagttt gtttaaggac aaaaaattag tagaatactt aaaggagagt     420
tgttttgatg cagtgtttct cgatcctttt gatgcctgtg gcttaattgt tgccaaatat     480
ttctccctcc cctctgtggt cttcgccagg ggaatatttt gccactatct tgaagaaggt     540
gcacagtgcc ctgctcctct ttcctatgtc cccagacttc tcttagggtt ctcagacgcc     600
atgactttca aggagagagt atggaaccac atcatgcact tggaggaaca tttattttgc     660
ccctattttt tcaaaaatgt cttagaaata gcctctgaaa ttctccaaac ccctgtcacg     720
gcatatgatc tctacagcca cacatcaatt tggttgttgc gaactgactt tgttttggag     780
tatcccaaac ccgtgatgcc caatatgatc ttcattggtg gtatcaactg tcatcaggga     840
aagccagtgc ctatggtaag ttatctcccc tttagcacat taagaataat ctggctttgg     900
aaattaaaag atttcttaca gaatcataat ttatcattta catttgtccc atttggaatt     960
tctttctggt ttaaggaatt cttttgtacc aattcactta attgttgggt agcaaattgt    1020
ataaagcagc tcttgttgat atgtaagtgt atacaattga tataattgta gatcatatct    1080
aggctgcaat ctaaatgcta tttttggaaa aatacaaaaa aaccacagta agaaatgaaa    1140
cttccctttt tttgctaatt ctacactacc cccagaggaa aatattctta gcagttttgt    1200
gtgaattgtt ttcaattttt ttgaaatta                                      1229
```

<210> 69
<211> 530
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (1)...(530)
<223> UGT1A9*1 protein

<400> 69

```
Met Ala Cys Thr Gly Trp Thr Ser Pro Leu Pro Leu Cys Val Cys Leu
 1               5                  10                  15
Leu Leu Thr Cys Gly Phe Ala Glu Ala Gly Lys Leu Leu Val Val Pro
            20                  25                  30
Met Asp Gly Ser His Trp Phe Thr Met Arg Ser Val Val Glu Lys Leu
        35                  40                  45
Ile Leu Arg Gly His Glu Val Val Val Val Met Pro Glu Val Ser Trp
    50                  55                  60
```

Gln Leu Gly Arg Ser Leu Asn Cys Thr Val Lys Thr Tyr Ser Thr Ser
65 70 75 80
Tyr Thr Leu Glu Asp Leu Asp Arg Glu Phe Lys Ala Phe Ala His Ala
85 90 95
Gln Trp Lys Ala Gln Val Arg Ser Ile Tyr Ser Leu Leu Met Gly Ser
100 105 110
Tyr Asn Asp Ile Phe Asp Leu Phe Phe Ser Asn Cys Arg Ser Leu Phe
115 120 125
Lys Asp Lys Lys Leu Val Glu Tyr Leu Lys Glu Ser Ser Phe Asp Ala
130 135 140
Val Phe Leu Asp Pro Phe Asp Asn Cys Gly Leu Ile Val Ala Lys Tyr
145 150 155 160
Phe Ser Leu Pro Ser Val Val Phe Ala Arg Gly Ile Leu Cys His Tyr
165 170 175
Leu Glu Glu Gly Ala Gln Cys Pro Ala Pro Leu Ser Tyr Val Pro Arg
180 185 190
Ile Leu Leu Gly Phe Ser Asp Ala Met Thr Phe Lys Glu Arg Val Arg
195 200 205
Asn His Ile Met His Leu Glu Glu His Leu Leu Cys His Arg Phe Phe
210 215 220
Lys Asn Ala Leu Glu Ile Ala Ser Glu Ile Leu Gln Thr Pro Val Thr
225 230 235 240
Glu Tyr Asp Leu Tyr Ser His Thr Ser Ile Trp Leu Leu Arg Thr Asp
245 250 255
Phe Val Leu Asp Tyr Pro Lys Pro Val Met Pro Asn Met Ile Phe Ile
260 265 270
Gly Gly Ile Asn Cys His Gln Gly Lys Pro Leu Pro Met Glu Phe Glu
275 280 285
Ala Tyr Ile Asn Ala Ser Gly Glu His Gly Ile Val Val Phe Ser Leu
290 295 300
Gly Ser Met Val Ser Glu Ile Pro Glu Lys Lys Ala Met Ala Ile Ala
305 310 315 320
Asp Ala Leu Gly Lys Ile Pro Gln Thr Val Leu Trp Arg Tyr Thr Gly
325 330 335
Thr Arg Pro Ser Asn Leu Ala Asn Asn Thr Ile Leu Val Lys Trp Leu
340 345 350
Pro Gln Asn Asp Leu Leu Gly His Pro Met Thr Arg Ala Phe Ile Thr
355 360 365
His Ala Gly Ser His Gly Val Tyr Glu Ser Ile Cys Asn Gly Val Pro
370 375 380
Met Val Met Met Pro Leu Phe Gly Asp Gln Met Asp Asn Ala Lys Arg
385 390 395 400
Met Glu Thr Lys Gly Ala Gly Val Thr Leu Asn Val Leu Glu Met Thr
405 410 415
Ser Glu Asp Leu Glu Asn Ala Leu Lys Ala Val Ile Asn Asp Lys Ser
420 425 430
Tyr Lys Glu Asn Ile Met Arg Leu Ser Ser Leu His Lys Asp Arg Pro
435 440 445
Val Glu Pro Leu Asp Leu Ala Val Phe Trp Val Glu Phe Val Met Arg
450 455 460
His Lys Gly Ala Pro His Leu Arg Pro Ala Ala His Asp Leu Thr Trp
465 470 475 480
Tyr Gln Tyr His Ser Leu Asp Val Ile Gly Phe Leu Leu Ala Val Val
485 490 495
Leu Thr Val Ala Phe Ile Thr Phe Lys Cys Cys Ala Tyr Gly Tyr Arg
500 505 510
Lys Cys Leu Gly Lys Lys Gly Arg Val Lys Lys Ala His Lys Ser Lys

515 520 525
Thr His
530

<210> 70
<211> 530
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (1)...(50)
<223> UGT1A9*2 protein

<400> 70
Met Ala Tyr Thr Gly Trp Thr Ser Pro Leu Pro Leu Cys Val Cys Leu
1 5 10 15
Leu Leu Thr Cys Gly Phe Ala Glu Ala Gly Lys Leu Leu Val Val Pro
20 25 30
Met Asp Gly Ser His Trp Phe Thr Met Arg Ser Val Val Glu Lys Leu
35 40 45
Ile Leu Arg Gly His Glu Val Val Val Val Met Pro Glu Val Ser Trp
50 55 60
Gln Leu Gly Arg Ser Leu Asn Cys Thr Val Lys Thr Tyr Ser Thr Ser
65 70 75 80
Tyr Thr Leu Glu Asp Leu Asp Arg Glu Phe Lys Ala Phe Ala His Ala
85 90 95
Gln Trp Lys Ala Gln Val Arg Ser Ile Tyr Ser Leu Leu Met Gly Ser
100 105 110
Tyr Asn Asp Ile Phe Asp Leu Phe Phe Ser Asn Cys Arg Ser Leu Phe
115 120 125
Lys Asp Lys Lys Leu Val Glu Tyr Leu Lys Glu Ser Ser Phe Asp Ala
130 135 140
Val Phe Leu Asp Pro Phe Asp Asn Cys Gly Leu Ile Val Ala Lys Tyr
145 150 155 160
Phe Ser Leu Pro Ser Val Val Phe Ala Arg Gly Ile Leu Cys His Tyr
165 170 175
Leu Glu Glu Gly Ala Gln Cys Pro Ala Pro Leu Ser Tyr Val Pro Arg
180 185 190
Ile Leu Leu Gly Phe Ser Asp Ala Met Thr Phe Lys Glu Arg Val Arg
195 200 205
Asn His Ile Met His Leu Glu Glu His Leu Leu Cys His Arg Phe Phe
210 215 220
Lys Asn Ala Leu Glu Ile Ala Ser Glu Ile Leu Gln Thr Pro Val Thr
225 230 235 240
Glu Tyr Asp Leu Tyr Ser His Thr Ser Ile Trp Leu Leu Arg Thr Asp
245 250 255
Phe Val Leu Asp Tyr Pro Lys Pro Val Met Pro Asn Met Ile Phe Ile
260 265 270
Gly Gly Ile Asn Cys His Gln Gly Lys Pro Leu Pro Met Glu Phe Glu
275 280 285
Ala Tyr Ile Asn Ala Ser Gly Glu His Gly Ile Val Val Phe Ser Leu
290 295 300
Gly Ser Met Val Ser Glu Ile Pro Glu Lys Lys Ala Met Ala Ile Ala
305 310 315 320
Asp Ala Leu Gly Lys Ile Pro Gln Thr Val Leu Trp Arg Tyr Thr Gly
325 330 335

Thr Arg Pro Ser Asn Leu Ala Asn Asn Thr Ile Leu Val Lys Trp Leu
340 345 350
Pro Gln Asn Asp Leu Leu Gly His Pro Met Thr Arg Ala Phe Ile Thr
355 360 365
His Ala Gly Ser His Gly Val Tyr Glu Ser Ile Cys Asn Gly Val Pro
370 375 380
Met Val Met Met Pro Leu Phe Gly Asp Gln Met Asp Asn Ala Lys Arg
385 390 395 400
Met Glu Thr Lys Gly Ala Gly Val Thr Leu Asn Val Leu Glu Met Thr
405 410 415
Ser Glu Asp Leu Glu Asn Ala Leu Lys Ala Val Ile Asn Asp Lys Ser
420 425 430
Tyr Lys Glu Asn Ile Met Arg Leu Ser Ser Leu His Lys Asp Arg Pro
435 440 445
Val Glu Pro Leu Asp Leu Ala Val Phe Trp Val Glu Phe Val Met Arg
450 455 460
His Lys Gly Ala Pro His Leu Arg Pro Ala Ala His Asp Leu Thr Trp
465 470 475 480
Tyr Gln Tyr His Ser Leu Asp Val Ile Gly Phe Leu Leu Ala Val Val
485 490 495
Leu Thr Val Ala Phe Ile Thr Phe Lys Cys Cys Ala Tyr Gly Tyr Arg
500 505 510
Lys Cys Leu Gly Lys Lys Gly Arg Val Lys Lys Ala His Lys Ser Lys
515 520 525
Thr His
530

<210> 71
<211> 530
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (1)...(530)
<223> UGT1A9*3 protein

<400> 71
Met Ala Cys Thr Gly Trp Thr Ser Pro Leu Pro Leu Cys Val Cys Leu
1 5 10 15
Leu Leu Thr Cys Gly Phe Ala Glu Ala Gly Lys Leu Leu Val Val Pro
20 25 30
Thr Asp Gly Ser His Trp Phe Thr Met Arg Ser Val Val Glu Lys Leu
35 40 45
Ile Leu Arg Gly His Glu Val Val Val Val Met Pro Glu Val Ser Trp
50 55 60
Gln Leu Gly Arg Ser Leu Asn Cys Thr Val Lys Thr Tyr Ser Thr Ser
65 70 75 80
Tyr Thr Leu Glu Asp Leu Asp Arg Glu Phe Lys Ala Phe Ala His Ala
85 90 95
Gln Trp Lys Ala Gln Val Arg Ser Ile Tyr Ser Leu Leu Met Gly Ser
100 105 110
Tyr Asn Asp Ile Phe Asp Leu Phe Phe Ser Asn Cys Arg Ser Leu Phe
115 120 125
Lys Asp Lys Lys Leu Val Glu Tyr Leu Lys Glu Ser Ser Phe Asp Ala
130 135 140
Val Phe Leu Asp Pro Phe Asp Asn Cys Gly Leu Ile Val Ala Lys Tyr

```
145                 150                 155                 160
Phe Ser Leu Pro Ser Val Val Phe Ala Arg Gly Ile Leu Cys His Tyr
                165                 170                 175
Leu Glu Glu Gly Ala Gln Cys Pro Ala Pro Leu Ser Tyr Val Pro Arg
            180                 185                 190
Ile Leu Leu Gly Phe Ser Asp Ala Met Thr Phe Lys Glu Arg Val Arg
        195                 200                 205
Asn His Ile Met His Leu Glu Glu His Leu Leu Cys His Arg Phe Phe
    210                 215                 220
Lys Asn Ala Leu Glu Ile Ala Ser Glu Ile Leu Gln Thr Pro Val Thr
225                 230                 235                 240
Glu Tyr Asp Leu Tyr Ser His Thr Ser Ile Trp Leu Leu Arg Thr Asp
                245                 250                 255
Phe Val Leu Asp Tyr Pro Lys Pro Val Met Pro Asn Met Ile Phe Ile
            260                 265                 270
Gly Gly Ile Asn Cys His Gln Gly Lys Pro Leu Pro Met Glu Phe Glu
        275                 280                 285
Ala Tyr Ile Asn Ala Ser Gly Glu His Gly Ile Val Val Phe Ser Leu
    290                 295                 300
Gly Ser Met Val Ser Glu Ile Pro Glu Lys Lys Ala Met Ala Ile Ala
305                 310                 315                 320
Asp Ala Leu Gly Lys Ile Pro Gln Thr Val Leu Trp Arg Tyr Thr Gly
                325                 330                 335
Thr Arg Pro Ser Asn Leu Ala Asn Asn Thr Ile Leu Val Lys Trp Leu
            340                 345                 350
Pro Gln Asn Asp Leu Leu Gly His Pro Met Thr Arg Ala Phe Ile Thr
        355                 360                 365
His Ala Gly Ser His Gly Val Tyr Glu Ser Ile Cys Asn Gly Val Pro
    370                 375                 380
Met Val Met Met Pro Leu Phe Gly Asp Gln Met Asp Asn Ala Lys Arg
385                 390                 395                 400
Met Glu Thr Lys Gly Ala Gly Val Thr Leu Asn Val Leu Glu Met Thr
                405                 410                 415
Ser Glu Asp Leu Glu Asn Ala Leu Lys Ala Val Ile Asn Asp Lys Ser
            420                 425                 430
Tyr Lys Glu Asn Ile Met Arg Leu Ser Ser Leu His Lys Asp Arg Pro
        435                 440                 445
Val Glu Pro Leu Asp Leu Ala Val Phe Trp Val Glu Phe Val Met Arg
    450                 455                 460
His Lys Gly Ala Pro His Leu Arg Pro Ala Ala His Asp Leu Thr Trp
465                 470                 475                 480
Tyr Gln Tyr His Ser Leu Asp Val Ile Gly Phe Leu Leu Ala Val Val
                485                 490                 495
Leu Thr Val Ala Phe Ile Thr Phe Lys Cys Cys Ala Tyr Gly Tyr Arg
            500                 505                 510
Lys Cys Leu Gly Lys Lys Gly Arg Val Lys Lys Ala His Lys Ser Lys
        515                 520                 525
Thr His
    530
```

THIS PAGE BLANK (USPTO)

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☑ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ______________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)